

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 18.1432 Seconds  
(without alignments)  
636.214 Million cell updates/sec

Title: US-10-036-444-4

Perfect score: 632

Sequence: 1 LWSQPPPIRTLEGSAPFLP.....TGNGTRLVVEKEHPQLGAGT 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283365

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	18.4	132	2 A24402	T-cell receptor al
2	114.5	18.1	136	2 B45893	T-cell receptor al
3	107.5	17.0	138	2 C27577	T-cell receptor al
4	105	16.6	131	2 F45893	T-cell receptor al
5	103	16.3	129	2 A42692	T-cell receptor al
6	102.5	16.2	110	2 B24092	T-cell receptor al
7	101	16.0	120	2 I54487	T-cell receptor al
8	99.5	15.7	131	2 E24092	T-cell receptor al
9	99.5	15.7	131	2 D24092	T-cell receptor al
10	99	15.7	132	1 RWSAV	T-cell receptor al
11	98.5	15.6	130	2 A31211	T-cell receptor al
12	98.5	15.6	146	2 S26408	T-cell receptor be
13	97	15.3	139	2 S36325	T-cell receptor de
14	95.5	15.1	218	2 B47712	myelin/oligodendro
15	95.5	15.1	271	2 A53268	T-cell receptor al
16	94	14.9	526	2 S70587	butyrophilin precu
17	93.5	14.8	247	2 A57117	myelin/oligodendro
18	93	14.7	110	2 A24092	T-cell receptor al
19	93	14.7	117	2 I68824	T-cell receptor al
20	93	14.7	1694	2 S50065	sialoadhesin - mou
21	92.5	14.6	132	2 D45893	T-cell receptor al
22	92.5	14.6	139	2 S36302	T-cell receptor de
23	92.5	14.6	267	1 RWSVC8	T-cell receptor al
24	91.5	14.5	107	2 S60590	IG heavy chain var
25	90	14.2	132	2 A24092	T-cell receptor al
26	89	14.1	134	2 A45893	T-cell receptor al
27	88.5	14.0	137	2 C45893	T-cell receptor al
28	88.5	14.0	223	2 A29063	cytotoxic T-lympho
29	88.5	14.0	264	2 F27579	T-cell receptor al

30 88.5 14.0 1197 2 T30581 neural cell adhesi  
31 88 13.9 131 2 D24402 T-cell receptor al  
32 88 13.9 226 2 A45477 membrane-bound imm  
33 88 13.9 946 1 A47299 tor-related recept  
34 87.5 13.8 111 2 JH0333 T-cell receptor al  
35 87.5 13.8 132 2 S23374 T-cell receptor al  
36 87.5 13.8 246 2 A47712 myelin/oligodendro  
37 87.5 13.8 247 2 S58394 T-cell receptor al  
38 87 13.8 108 2 JH0342 T-cell receptor al  
39 86.5 13.7 95 2 G30603 T-cell receptor al  
40 86.5 13.7 129 2 S03478 hypothetical prote  
41 86.5 13.7 874 2 T29548 sax-3 protein - Ca  
42 86.5 13.7 1273 2 T42405 ig lambda chain pr  
43 86 13.6 140 2 PH0134 T-cell receptor al  
44 85.5 13.5 134 2 C29774 T-cell receptor al  
45 85.5 13.5 223 2 T09536 cytotoxic T-lympho

## ALIGNMENTS

### RESULT 1

A24402

T-cell receptor alpha chain precursor V region (CS) - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 30-May-1997

C/Accession: A24402

R/Becker, D.M.; Patten, P.; Chien, Y.; Yokota, T.; Eshhar, Z.; Giedlin, M.; Gascoigne, J.

Nature 317, 430-434, 1985

A/Title: Variability and repertoire size of T-cell receptor V-alpha gene segments.

A/Reference number: A93368, UID:86014379; PMID:2995827

A/Accession: A24402

A/Molecule type: mRNA

A/Residues: 1-132 <BEC>

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: T-cell receptor

Query Match 18.4%; Score 116; DB 2; Length 132;

Best Local Similarity 33.3%; Pred. No. 0.00019;

Matches 41; Conservative 16; Mismatches 42; Indels 24; Gaps 8;

Qy 3 VSQPPPIRTLEGSAPFLP...TGNGTRLVVEKEHPQLGAGT---NGTPE 54

Db 23 VQSPESLIVPEGAMVSLNCSFSDS---ASQSIWYQQH--PGKPKALISIFSNK 76

Qy 55 FROGLAPLASSRPLHDOELHVRGHDASIVCRVEVLGVGT---GNGTRLVVEK 111

Db 77 -EGRLTVYLNRLASLH---VSLHIKQSPSDSAVYLCAVRRSGANTGKLTFGHGTILRV-- 130

Qy 112 EHP 114

Db 131 -HP 132

### RESULT 2

B45893

T-cell receptor alpha chain precursor V region (BTA29) - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000

C/Accession: B45893

R/Ishiguro, N.; Tanaka, A.; Shinagawa, M.

Immunogenetics 31, 57-60, 1990

A/Title: Sequence analysis of bovine T-cell receptor alpha chain.

A/Reference number: A45893, UID:90129157; PMID:2137108

A/Accession: B45893

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-136 <ISH>

A/Cross-references: GB:D90011; NID:G217610; PIDN:BA414061.1; PID:G217611

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: T-cell receptor

Query Match 18.1%; Score 114.5; DB 2; Length 136;

Best Local Similarity 29.5%; Pred. No. 0.00028;  
Matches 36; Conservative 19; Mismatches 39; Indels 29; Gaps 7;  
QY 4 SQPEPTLGGSAFLPCSFNASQGRLAIGSVTWFRDEVPVPGKEVNGTPER-----55  
Db 27 AEPASLPVPEGAASLGCTYSNSLY---FWYRQ--YPGK-----GPEFLQVYANN 75  
QY 56 ---RGRLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGT-----GNGTRLV 108  
Db 76 NKEGKFT--AQNKNKH-VSLIRDSPSDSATVLCVADTISTAGTKLTFGEGTRLI 132  
QY 109 VE 110  
Db 133 VK 134  
RESULT 3  
T-cell receptor alpha chain V region (5/10-20D) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 30-May-1997  
C:Accession: C27577; C27557  
R:Iwamoto, A.; Ohashi, P.S.; Pիրcher, H.; Walker, C.L.; Michalopoulos, E.E.; Rupp, F.; H  
J. Exp. Med. 165, 591-600, 1987  
A:Title: T cell receptor variable gene usage in a specific cytotoxic T cell response. P  
A:Reference number: A27557; MUID:87139812; PMID:3493320  
A:Accession: C27577  
A:Molecule type: mRNA  
A:Residues: 1-138 <IWA>  
A:Accession: C27557  
A:Molecule type: mRNA  
A:Residues: 22-138 <TW2>  
A:Cross-references: EMBL:X05733  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor  
Query Match 17.0%; Score 107.5; DB 2; Length 138;  
Best Local Similarity 13.9%; Pred. No. 0.0014;  
Matches 38; Conservative 17; Mismatches 45; Indels 19; Gaps 7;  
QY 3 VSQPE-IRTEGSAFLPCSFNASQGRLAIGSVTWFRDEVPVPGKEV-----NGTPE 54  
Db 24 VQSPESLIIVPEGAMTSLNCTFSDASQY----FAWYRQH--SGKAPKALMSIFSNGEKE 77  
QY 55 FRGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVL-CLGVGTGNGTRLVVEKE 112  
Db 78 -EGFTTHNKASLH---FSLHIRDSPSDSATVLCVADTISTAGTKLTFGEGTRLVSPD 132  
RESULT 4  
T-cell receptor alpha chain precursor V region (BTA25) - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000  
C:Accession: F45893  
R:Ishiguro, N.; Tanaka, A.; Shinagawa, M.  
Immunogenetics 31, 57-60, 1990  
A:Title: Sequence analysis of bovine T-cell receptor alpha chain.  
A:Reference number: A45893; MUID:90129157; PMID:2137108  
A:Accession: F45893  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-131 <ISH>  
A:Cross-references: GB:D90015; NID:G217618; PIDN:BAAL4065.1; PID:G217619  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor  
F:36-109/Domain: immunoglobulin homology <IMM>  
Query Match 16.6%; Score 105; DB 2; Length 131;  
Best Local Similarity 30.3%; Pred. No. 0.0024;  
Matches 36; Conservative 23; Mismatches 34; Indels 26; Gaps 8;  
QY 3 VSQPEIRTL-EGSSAFLPCSFNASQGRLAIGSVTWFRDEVPVPGKE-----VRNGTPEF 55

Db 24 VEQSPVLSQEGANSTLRCNFSDT-----VDSVQWFQON--PGGALTTLFFIASGTKK- 75  
QY 56 RGRLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGT-----GNGTRLV 109  
Db 76 NERMSSTVNSK---ERYSTLHTASQLEDAATYLCVAD---LGSNGRLVFGKGTSLAV 128  
RESULT 5  
A42692  
T-cell receptor alpha chain (lysozyme peptide specific) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 23-Jul-1999  
C:Accession: A42692  
R:Kobori, J.A.; Hood, L.; Shastri, N.  
Proc. Natl. Acad. Sci. U.S.A. 89, 2940-2944, 1992  
A:Title: Structure-function relationship among T-cell receptors specific for lysozyme p  
A:Reference number: A42692; MUID:92212944; PMID:1313573  
A:Accession: A42692  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-129 <KOB>  
A:Cross-references: GB:M87844; NID:G201804; PIDN:AAAS1236.1; PID:G554375  
A:Experimental source: T-cell hybrid BO4H.9.1  
A:Note: sequence extracted from NCBI backbone (NCBIN:92850, NCBI:P:92851)  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor  
F:36-109/Domain: immunoglobulin homology <IMM>  
Query Match 16.3%; Score 103; DB 2; Length 129;  
Best Local Similarity 29.8%; Pred. No. 0.0037;  
Matches 39; Conservative 14; Mismatches 34; Indels 44; Gaps 8;  
QY 2 WVS-----QPPEIRTEGSAFLPCSFNASQGRLAIGSVTWFRDE-----VPG 45  
Db 18 WVSQDKVKQSPSALSQEGTNSALRCNFS-----IAATTVQVFLQNGRSLINLFYLP- 71  
QY 46 KEVNGTPPEPRGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTGN-- 103  
Db 72 -----GTYE-NGRLKSAFDSK---ESYSTLHIDAQLEDSGTFFCAA- 116  
QY 104 -----GTRLVV 109  
Db 117 YVFGAGTRLKV 127  
RESULT 6  
B24092  
T-cell receptor alpha chain V region (5C.C7) - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 21-Jan-2000  
C:Accession: B24092  
R:Fink, P.J.; Matig, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.  
Nature 321, 219-226, 1986  
A:Title: Correlations between T-cell specificity and the structure of the antigen recep  
A:Reference number: A93380; MUID:86230843; PMID:3012351  
A:Accession: B24092  
A:Molecule type: mRNA  
A:Residues: 1-110 <FTN>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor  
F:16-89/Domain: immunoglobulin homology <IMM>  
Query Match 16.2%; Score 102.5; DB 2; Length 110;  
Best Local Similarity 31.9%; Pred. No. 0.0035;  
Matches 36; Conservative 16; Mismatches 46; Indels 15; Gaps 6;  
QY 3 VSQPEIRTL-EGSSAFLPCSFNASQGRLAIGSVTW- ----RDEVVPGKEVNGTPEFRG 57  
Db 4 VEQSPSALSLEHGSGSALRCNFTT-----MRAVQWFRKNSRGLINLFLASGTKE-NG 57  
QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLG-VGTGNGTRLVV 109



Db 58 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAEASNTKVVFGTGRLOV 107

## RESULT 7

I54487  
T-cell receptor alpha chain - mouse (fragment)  
C/Species: Mus musculus (house mouse)  
C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 21-Jan-2000  
C/Accession: I54487; I58823  
R/Nakajima, F.B.; Di Vincenzo, J.P.; Jameson, S.C.; Gascoigne, N.R.J.  
Immunogenetics 35, 190-198, 1992  
A/Title: Chromosome 14 in B10.A (18R) mice is recombinant and includes Tcra-V alleles.  
A/Reference number: I54487; MUID:92165347; PMID:1371499  
A/Accession: I54487  
A/Status: preliminary; translated from GB/EMBL/DBBJ  
A/Molecule type: mRNA  
A/Residues: 1-120 <RES>  
A/Cross-references: GB:M55634; NID:gl99677; PIDN:AAA39701.1; PID:gl99678  
A/Accession: I68823  
A/Status: preliminary; translated from GB/EMBL/DBBJ  
A/Molecule type: mRNA  
A/Residues: 104-120 <RES2>  
A/Cross-references: GB:M55634; NID:gl99677; PIDN:AAA39702.1; PID:gl99679  
A/Note: J-alpha TA61  
C/Genetics:  
A/Gene: MHC-V-alpha-11; J-alpha-TA61  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: transmembrane protein  
F/26-100/Domain: immunoglobulin homology <IMM>

Query Match 16.0%; Score 101; DB 2; Length 120;  
Best Local Similarity 31.6%; Pred. No. 0.0054;  
Matches 36; Conservative 17; Mismatches 47; Indels 14; Gaps 6;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFRD---EVVPGKEVRNGTPEFRG 57  
DB 14 VEQSPALSLSHEGTGSALRCNFTTT-----MRSVQWFRQNSRGSLSILFYLASGTKE-NG 57

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLVGVTGNTRLVWEKEHPQLG 110  
DB 68 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAADTNTGKLTFGDGTVLTVK 119

## RESULT 8

E24092  
T-cell receptor alpha chain precursor V region (4.C3) - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 19-Nov-1988 #sequence\_revision 19-Nov-1988 #text\_change 21-Jan-2000  
C/Accession: E24092  
R/Fink, P.J.; Matis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.  
Nature 321, 219-226, 1986  
A/Title: Correlations between T-cell specificity and the structure of the antigen receptor  
A/Reference number: A93380; MUID:86230843; PMID:3012351  
A/Accession: E24092  
A/Molecule type: mRNA  
A/Residues: 1-131 <FIN>  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: T-cell receptor  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-131/Product: T-cell receptor alpha chain V region 4.C3 #status predicted <MAT>  
F/36-109/Domain: immunoglobulin homology <IMM>

Query Match 15.7%; Score 99.5; DB 2; Length 131;  
Best Local Similarity 29.3%; Pred. No. 0.0085;  
Matches 36; Conservative 18; Mismatches 42; Indels 27; Gaps 7;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFP----RDEVVPGKEVRNGTPEFRG 57  
DB 24 VEQSPALSLSHEGTGSALRCNFTTT-----MRAVQWFRKNSRGSLSILFYLASGTKE-NG 77

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLVGVTGNTRLVWEKEHPQLG 117  
DB 78 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAE-----STSSQKLV-----FG 121

QY 118 AGT 120  
DB 122 QGT 124

## RESULT 9

E24092  
T-cell receptor alpha chain precursor V region (B10) - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 19-Nov-1988 #sequence\_revision 01-Dec-2000 #text\_change 11-Jan-2002  
C/Accession: E24092; S03507  
R/Fink, P.J.; Matis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.  
Nature 321, 219-226, 1986  
A/Title: Correlations between T-cell specificity and the structure of the antigen receptor  
A/Reference number: A93380; MUID:86230843; PMID:3012351  
A/Accession: E24092  
A/Molecule type: mRNA  
A/Residues: 1-131 <FIN>  
R/Winoto, A.; Mjolsness, S.; Hood, L.  
Nature 316, 832-836, 1985  
A/Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.  
A/Reference number: S03503; MUID:85296332; PMID:2993908  
A/Accession: S03507  
A/Molecule type: DNA  
A/Residues: 111-130 <WIN>  
A/Cross-references: EMBL:X03057; NID:954519; PIDN:CAA26864.1; PID:91334131  
A/Note: this sequence was determined from the germline gene  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: T-cell receptor  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-131/Product: T-cell receptor alpha chain V region B10 #status predicted <MAT>  
F/36-109/Domain: immunoglobulin homology <IMM>

Query Match 15.7%; Score 99.5; DB 2; Length 131;  
Best Local Similarity 29.3%; Pred. No. 0.0085;  
Matches 36; Conservative 18; Mismatches 42; Indels 27; Gaps 7;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFP----RDEVVPGKEVRNGTPEFRG 57  
DB 24 VEQSPALSLSHEGTGSALRCNFTTT-----MRAVQWFRKNSRGSLSILFYLASGTKE-NG 77

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLVGVTGNTRLVWEKEHPQLG 117  
DB 78 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAE-----ATSSQKLV-----FG 121

QY 118 AGT 120  
DB 122 QGT 124

## RESULT 10

RWMSAV  
T-cell receptor alpha chain precursor V region (2B4) - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 30-May-1997  
C/Accession: A02015  
R/Chien, Y.; Becker, D.M.; Lindsten, T.; Okamura, M.; Cohen, D.I.; Davis, M.M.  
Nature 312, 31-35, 1984  
A/Title: A third type of murine T-cell receptor gene.  
A/Reference number: A93344; MUID:85036634; PMID:6548551  
A/Accession: A02015  
A/Molecule type: mRNA  
A/Residues: 1-132 <CHI>  
A/Experimental source: hybridoma 2B4, clone Tt11  
C/Superfamily: immunoglobulin V region; immunoglobulin homology  
C/Keywords: glycoprotein; heterotetramer; receptor; T-cell  
F/1-20/Domain: signal sequence #status predicted <SIG>  
F/21-132/Product: T-cell receptor alpha chain V region (2B4) #status predicted <MAT>  
F/21-113/Region: V segment  
F/114-117/Region: D segment  
F/118-132/Region: J segment  
F/42/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.7%; Score 99.5; DB 2; Length 131;  
Best Local Similarity 29.3%; Pred. No. 0.0085;  
Matches 36; Conservative 18; Mismatches 42; Indels 27; Gaps 7;

QY 3 VSQPPERTL-EGSSAFPCSFNASQGRLAIGSVTFP----RDEVVPGKEVRNGTPEFRG 57  
DB 24 VEQSPALSLSHEGTGSALRCNFTTT-----MRAVQWFRKNSRGSLSILFYLASGTKE-NG 77

QY 58 RLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLVGVTGNTRLVWEKEHPQLG 117  
DB 78 RLKSAFDSK---ERYSTLHIRDAQLSDSGTYFCAAE-----STSSQKLV-----FG 121



Search completed: February 26, 2004, 12:14:50  
Job time : 19.1432 secs

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 10.1857 Seconds  
(without alignments)  
613.452 Million cell updates/sec

Title: US-10-036-444-4  
Perfect score: 632  
Sequence: 1 LWSQPPEIRTEGSSAFPLP.....TGNTRLVVEKEHPQLGAGT 120

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	15.7	132	1	TVA2 MOUSE
2	95.5	15.1	245	1	MOG_RAT
3	94	14.9	526	1	BUTY HUMAN
4	93.5	14.8	246	1	MOG MOUSE
5	93	14.7	1694	1	SN MOUSE
6	88.5	14.0	223	1	CTL4 MOUSE
7	88.5	14.0	1197	1	CAMI BRARE
8	88	13.9	226	1	C79A HUMAN
9	87.5	13.8	524	1	BUTY MOUSE
10	87.5	13.8	223	1	CTL4 HUMAN
11	87.5	13.8	246	1	MOG BOVIN
12	87.5	13.8	247	1	MOG HUMAN
13	85.5	13.5	223	1	CTL4 PIG
14	84.5	13.4	223	1	CTL4 RABIT
15	83	13.1	246	1	MYPO HETFR
16	82.5	13.1	403	1	CD33 MOUSE
17	82	13.0	467	1	SIL5 MOUSE
18	82	13.0	3707	1	PGBM MOUSE
19	81.5	12.9	646	1	MUI8 HUMAN
20	81	12.8	526	1	BUTY BOVIN
21	79	12.5	551	1	SIL5 HUMAN
22	78.5	12.4	111	1	LV2H HUMAN
23	78.5	12.4	463	1	SIL9 HUMAN
24	77	12.2	134	1	TVB7 MOUSE
25	77	12.2	270	1	BASI RABIT
26	77	12.2	385	1	AMPC LYSLA
27	76.5	12.1	365	1	CMAR HUMAN
28	76	12.0	1051	1	PTK7 CHICK
29	75	11.9	365	1	CMAR MOUSE
30	74.5	11.8	111	1	LV1D HUMAN
31	74.5	11.8	1452	1	PTPM HUMAN
32	74	11.7	215	1	CIB2 HUMAN
33	74	11.7	977	1	KEMS_MOUSE

34	74	11.7	978	1	KFMS RAT	Q00495	rattus norv
35	74	11.7	1493	1	NEO1_MOUSE	P97798	mus musculu
36	73.5	11.6	106	1	LV4E HUMAN	P01716	homo sapien
37	73.5	11.6	108	1	LV5A HUMAN	P01719	homo sapien
38	73.5	11.6	111	1	LV2D HUMAN	P01707	homo sapien
39	73.5	11.6	111	1	LV2H HUMAN	P01711	homo sapien
40	73.5	11.6	438	1	KMLS SHEEP	O02827	ovis aries
41	73.5	11.6	2132	1	PGCA MOUSE	O61282	mus musculu
42	73	11.6	110	1	KV15 RABIT	P01696	oryctolagus
43	73	11.6	130	1	TVAI_MOUSE	P01738	mus musculu
44	73	11.6	898	1	PAS2_SCHAM	P22648	schistocerc
45	73	11.6	1259	1	CAML_RAT	Q05695	rattus norv

## ALIGNMENTS

RESULT 1																								
TVA2_MOUSE																								
ID	TVA2_MOUSE	STANDARD;	PRT;	132	AA.																			
AC	P01739;																							
DT	21-JUL-1986 (Rel. 01, Created)																							
DT	21-JUL-1986 (Rel. 01, Last sequence update)																							
DT	10-OCT-2003 (Rel. 42, Last annotation update)																							
DE	T-cell receptor alpha chain V region 2B4 precursor.																							
OS	Mus musculus (Mouse).																							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;																							
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.																							
OX	NCBI_TaxID=10090;																							
RN	[1]																							
RP	SEQUENCE FROM N.A. (CLONE TT11).																							
RX	MEDLINE=85036634; PubMed=6548551;																							
RA	Chien Y., Becker D.M., Lindsten T., Okamura M., Cohen D.I.,																							
RA	Davis M.M.;																							
RT	"A third type of murine T-cell receptor gene.";																							
RL	Nature 312:31-35(1984).																							
DR	PIR; A02015; RMSAV.																							
DR	HSSP; P01607; 1REI.																							
DR	InterPro; IPR007110; Ig-like.																							
DR	InterPro; IPR003596; Ig_V.																							
DR	Pfam; PF00047; Ig; 1.																							
DR	SMART; SM00406; IGV; 1.																							
DR	PROSITE; PS0835; IG LIKE; 1.																							
KW	T-cell; Receptor; Immunoglobulin domain; Glycoprotein; Signal.																							
FT	SIGNAL	1	20																					
FT	CHAIN	21	132																					
FT	DOMAIN	21	113																					
FT	DOMAIN	114	117																					
FT	DOMAIN	118	132																					
FT	CARBOHYD	42	42																					
FT	NON_TER	132	132																					
SQ	SEQUENCE	132	AA;	14668	MW;	CABBF6CF1DD3448B	CRC64;																	
Query Match																								
		15.7%;		Score 99;		DB 1;		Length 132;																
		Best Local Similarity		29.4%;		Pred. No. 0.0025;																		
		Matches 37;		Conservative 17;		Mismatches 42;		Indels 30;																
								Gaps 8;																
QY	3	VSQPPE-IRTEGSSAFPLCSFNASQCRALGISTVTFERDEVPGKEVR	-----NGTPE	54																				
DB	24	VQQPESLIIVEGARTSLNCTFSASQY	-----FWYRQH--SGKAPKALMSIFSNGEKE	77																				
QY	55	FRGLAPLASRFLHDHQAELHIRDVGHDAISIVVCVRVELGLGVGTGNGTRLVVEKEHP	114																					
DB	78	EGRFTIHLNKASLH---	FSLHIRDSPSDSALYLCATLYG-----GSGNKLI-----	122																				
QY	115	QLGAGT	120																					
DB	123	-FGTGT	127																					
RESULT 2																								
MOG_RAT																								
ID	MOG_RAT	STANDARD;	PRT;	245	AA.																			

AC Q63345;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Myelin-oligodendrocyte glycoprotein precursor.  
 GN MOG.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=93085763; PubMed=1453482;  
 RA Gardiner M.V., Amiguet P., Lington C., Matthieu J.-M.;  
 RT "Myelin/oligodendrocyte glycoprotein is a unique member of the  
 RT immunoglobulin superfamily.";  
 RL J. Neurosci. Res. 33:177-187(1992).  
 RN [2]  
 RP SEQUENCE OF 28-245 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=93376728; PubMed=8367453;  
 RA Pham-Dinh D., Mattei M.-G., Nussbaum J.-L., Rousset G.,  
 RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.F.,  
 RA Dautigny A.;  
 RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the  
 RT immunoglobulin superfamily encoded within the major  
 RT histocompatibility complex.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).  
 RN [3]  
 RP STRUCTURE BY NMR OF 62-82  
 RX MEDLINE=97354172; PubMed=9210466;  
 RA Albores-Abo S., Wilson J.C., Bernard C.C.A., von Itzstein M.;  
 RT "A conformational study of the human and rat encephalitogenic myelin  
 RT oligodendrocyte glycoprotein peptides 35-55.";  
 RL Eur. J. Biochem. 246:59-70(1997).  
 CC -!- FUNCTION: Minor component of the myelin sheath. May be involved in  
 CC completion and/or maintenance of the myelin sheath and in cell-  
 CC cell communication.  
 CC -!- SUBUNIT: May form homodimers.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is  
 CC localized on the surface of myelin and oligodendrocyte cytoplasmic  
 CC membranes.  
 CC -!- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED  
 CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF  
 CC ACTIVE MYELINATION.  
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG  
 CC family.  
 CC -!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)  
 CC with oligodendrocyte-myelin glycoprotein (OMG).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M99485; AAA41628.1; -;  
 DR EMBL; L21995; AAF74786.1; -;  
 DR PIR; B47712; B47712.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003596; IG\_v.  
 DR Pfam; PF00047; IG\_1.  
 DR SMART; SM00406; IGv\_1.  
 DR PROSITE; PS0835; IG-LIKE; 1.  
 DR Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.  
 KW SIGNAL 1 27  
 FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.  
 FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 156 176 POTENTIAL.

FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 209 229 POTENTIAL.  
 FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 30 139 IG-LIKE.  
 FT DISULFID 51 125 POTENTIAL.  
 FT CARBOHYD 58 58 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 245 AA; 27881 MW; C97F8AD6D6A32B4 CRC64;  
 Query Match 15.1%; Score 95.5; DB 1; Length 245;  
 Best Local Similarity 36.2%; Pred. No. 0.011;  
 Matches 34; Conservative 8; Mismatches 33; Indels 19; Gaps 6;  
 QY 9 IRTLEGSSAFPCSFNASQRLAIG-SVTWFRDEVVPGKEV-----RNG-----TPEFR 56  
 DB 39 IRLVGDDELPC--RISPGKNTGMEVGYRS---PFSRVVHLRYRNGKQDAEAPEYR 93  
 QY 57 GLRLPLASSRFLHDHQAELHIRDVRGHDASIYVC 90  
 DB 94 GRTELLAKES--IGEKVALRIQNVRFSDGGYTC 125  
 RESULT 3  
 BUTY\_HUMAN  
 ID BUTY\_HUMAN STANDARD; PRT; 526 AA.  
 AC Q13410;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Butyrophilin precursor (BT) (Butyrophilin subfamily 1 member A1).  
 GN BTN1AL OR BTN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Breast;  
 RX MEDLINE=96201696; PubMed=8611614;  
 RA Taylor M.R., Peterson J.A., Ceriani R.L., Couto J.R.;  
 RT "Cloning and sequence analysis of human butyrophilin reveals a  
 RT potential receptor function.";  
 RL Biochim. Biophys. Acta 1306:1-4(1996).  
 CC -!- FUNCTION: May function in the secretion of milk-fat droplets. It  
 CC may act as a specific membrane-associated receptor for the  
 CC association of cytoplasmic droplets with the apical plasma  
 CC membrane (By similarity).  
 CC -!- SUBUNIT: Seems to associate with xanthine dehydrogenase/oxidase  
 CC (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG  
 CC family.  
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U39576; AAC50489.1; -;  
 DR EMBL; S70587; S70587.  
 DR Genew; HGNC:1135; BTN1AL.  
 DR MIM; 601610; -;  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0004872; F:receptor activity; TAS.  
 DR InterPro; IPR001870; B302.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003596; IG\_v.  
 DR InterPro; IPR006574; PRY.  
 DR InterPro; IPR003877; SPRY\_receptor.  
 DR Pfam; PF00047; IG\_1.



ID SN MOUSE STANDARD; PRT; 1694 AA.  
 AC Q62230; O55216; Q62228; Q62229;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE DE Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-  
 DE 1) (Sheep erythrocyte receptor) (SER).  
 DE 1) (Sheep erythrocyte receptor) (SER).  
 GN SN OR SA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=C57BL/6; TISSUE=Macrophage;  
 RX MEDLINE=95009950; PubMed=7925291;  
 RA Crocker P.R.; Mucklow S.; Boukron V.; McWilliam A.; Willis A.C.,  
 RA Gordon S.; Milon G.; Kelm S.; Bradfield P.;  
 RT "Sialoadhesin, a macrophage sialic acid binding receptor for  
 RT haemopoietic cells with 17 immunoglobulin-like domains.";  
 RL EMBO J. 13:4490-4503(1994).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=98051930; PubMed=9383289;  
 RX Mucklow S.; Gordon S.; Crocker P.R.;  
 RA "Characterization of the mouse sialoadhesin gene, Sn.";  
 RL Mamm. Genome 8:934-937(1997).  
 RL [3]  
 RN CHARACTERIZATION.  
 RC STRAIN=C57BL/6; TISSUE=Spleen;  
 RX MEDLINE=91266893; PubMed=2050106;  
 RA Crocker P.R.;  
 RT "Purification and properties of sialoadhesin, a sialic acid-binding  
 RT receptor of murine tissue macrophages.";  
 RL EMBO J. 10:1661-1669(1991).  
 RL [4]  
 RP SIALIC ACID BINDING.  
 RP MEDLINE=95179521; PubMed=7533044;  
 RA Kelm S.; Pelz A.; Schauer R.; Filbin M.T.; Tang S.; de Bellard M.E.,  
 RA Schnaar R.L.; Mahoney J.A.; Hartnell A.; Bradfield P.; Crocker P.R.;  
 RT "Sialoadhesin, myelin-associated glycoprotein and CD22 define a new  
 RT family of sialic acid-dependent adhesion molecules of the  
 RT immunoglobulin superfamily.";  
 RL Curr. Biol. 4:965-972(1994).  
 RL [5]  
 RN BINDING TO SPN.  
 RP MEDLINE=21136329; PubMed=11238599;  
 RA van den Berg T.K.; Nath D.; Ziltener H.J.; Vestweber D.; Fukuda M.,  
 RA van Die I.; Crocker P.R.;  
 RT "CD43 functions as a T cell counterreceptor for the macrophage  
 RT adhesion receptor sialoadhesin (Siglec-1).";  
 RL J. Immunol. 166:3637-3640(2001).  
 RL [6]  
 RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 20-138.  
 RP MEDLINE=98325385; PubMed=9660955;  
 RA May A.P.; Robinson R.C.; Vinson M.; Crocker P.R.; Jones E.Y.;  
 RT "Crystal structure of the N-terminal domain of sialoadhesin in  
 RT complex with 3' sialyllactose at 1.85 A resolution.";  
 RL Mol. Cell 1:719-728(1998).  
 RL [7]  
 RN STRUCTURE OF 20-138 BY NMR, AND MUTAGENESIS OF TRP-21 AND ARG-116.  
 RP MEDLINE=99321481; PubMed=10393093;  
 RA Crocker P.R.; Vinson M.; Kelm S.; Drickamer K.;  
 RT "Molecular analysis of sialoside binding to sialoadhesin by NMR and  
 RT site-directed mutagenesis.";  
 RL Biochem. J. 341:355-361(1999).  
 CC -!- FUNCTION: Macrophage-restricted adhesion molecule that mediates  
 CC sialic-acid dependent binding to lymphocytes, including  
 CC granulocytes, monocytes, natural killer cells, B-cells and CD8 T-  
 CC cells (By similarity). Preferentially binds to alpha2,3-linked  
 CC sialic acid. Binds to SPN/CD43 on T-cells. May play a role in  
 CC hemopoiesis.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and  
 CC soluble (isoforms 2 and 3).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=Q62230-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q62230-2; Sequence=VSP\_002573; VSP\_002574;  
 CC Name=3;  
 CC IsoId=Q62230-3; Sequence=VSP\_002575; VSP\_002576;  
 CC -!- TISSUE SPECIFICITY: Expressed by macrophages in various tissues.  
 CC High expression in spleen and lymph node with lower amounts in  
 CC lung, liver, bone marrow, heart and skin. No expression in thymus,  
 CC kidney, brain or small intestine.  
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC  
 CC (sialic acid binding Ig-like lectin) family.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -!- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; Z36293; CRA85290.1; -;  
 CC EMBL; Z36233; CRA85268.1; -;  
 CC EMBL; Z36234; CRA85269.1; -;  
 CC EMBL; U92842; AAB95641.1; -;  
 CC EMBL; U92833; AAB95641.1; JOINED.  
 CC EMBL; U92834; AAB95641.1; JOINED.  
 CC EMBL; U92836; AAB95641.1; JOINED.  
 CC EMBL; U92837; AAB95641.1; JOINED.  
 CC EMBL; U92838; AAB95641.1; JOINED.  
 CC EMBL; U92839; AAB95641.1; JOINED.  
 CC EMBL; U92840; AAB95641.1; JOINED.  
 CC EMBL; U92841; AAB95641.1; JOINED.  
 CC PIR; S50065; S50065.  
 CC PDB; 1QFO; 16-APR-99.  
 CC PDB; 1QFP; 16-APR-99.  
 CC MGD; MGI:99668; Sn.  
 CC GO; GO:0016021; C:integral to membrane; ISS.  
 CC GO; GO:0005529; F:sugar binding; ISS.  
 CC GO; GO:0016337; P:cell-cell adhesion; ISS.  
 CC GO; GO:0007160; P:cell-matrix adhesion; ISS.  
 CC GO; GO:0006954; P:inflammatory response; ISS.  
 CC InterPro; IPR007110; Ig\_LIKE.  
 CC InterPro; IPR003598; Ig\_C2.  
 CC InterPro; IPR003006; Ig\_MHC.  
 CC Pfam; PF00047; Ig\_16.  
 CC SMART; SM00408; IGC2; 5.  
 CC PROSITE; PS50835; IG\_LIKE; 14.  
 CC PROSITE; PS00290; IG\_MHC; 1.  
 CC Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;  
 CC Immunoglobulin domain; Repeat; Alternative splicing; 3D-structure.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1694 SIALOADHESIN.  
 FT DOMAIN 20 1638 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1639 1859 POTENTIAL.  
 FT DOMAIN 1660 1894 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 20 136 IG-LIKE V-TYPE.  
 FT DOMAIN 153 235 IG-LIKE C2-TYPE 1.  
 FT DOMAIN 239 321 IG-LIKE C2-TYPE 2.  
 FT DOMAIN 326 406 IG-LIKE C2-TYPE 3.  
 FT DOMAIN 416 508 IG-LIKE C2-TYPE 4.  
 FT DOMAIN 509 594 IG-LIKE C2-TYPE 5.  
 FT DOMAIN 602 701 IG-LIKE C2-TYPE 6.  
 FT DOMAIN 704 781 IG-LIKE C2-TYPE 7.  
 FT DOMAIN 795 890 IG-LIKE C2-TYPE 8.  
 FT DOMAIN 894 972 IG-LIKE C2-TYPE 9.  
 FT DOMAIN 979 1078 IG-LIKE C2-TYPE 10.  
 FT DOMAIN 1080 1160 IG-LIKE C2-TYPE 11.







```
FT CONFLICT 182 182 S -> T (IN REF. 2).
SQ SEQUENCE 223 AA; 24993 MW; 5318FAAF416F4685 CRC64;

Query Match
Best Local Similarity 14.0%; Score 88.5; DB 1; Length 223;
Matches 33; Conservative 17; Mismatches 58; Indels 13; Gaps 5;

QY 3 VSQPEIRTELGSSAFPCSFNAGSLAIGSVTWFRD-----EVVPGKEVNRGTPEPR 56
Db 40 VTQPSVVLASHGVASFPCSEYSHNTDEV-RVTLRQNDQKTEVCATFEKNTVGL 98
QY 57 GLRLAPLASSRFLDHQALHVRDGHDSIYVCRVEVL---GLGVGTGNGTRLVVEKEH 113
Db 99 D--YPCSGGF-NESRVNLFIQGLRAVDIGLYCKVELVPPPYFVGMGNGTQIVIDPE 155
QY 114 P 114
Db 156 P 156

CAML1_BRARE STANDARD; PRT; 1197 AA.
AC Q90478;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural cell adhesion molecule L1.1 (N-CAM L1.1) (Fragment).
GN NADL1.1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC TISSUE=Embryo;
RX MEDLINE=9615762; PubMed=8568941;
RA Tongiorgi E., Bernhardt R.R., Schachner M.;
RT "Zebrafish neurons express two L1-related molecules during early
axogenesis."
RL J. Neurosci. Res. 42:547-561(1995).
CC -1- FUNCTION: Cell adhesion molecule with an important role in the
development of the nervous system. Involved in neuron-neuron
adhesion, neurite fasciculation, outgrowth of neurites, etc. Binds
to axonin on neurons (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in postmitotic neurons in 16-36
hour embryos, including those in the brain, cranial ganglia and
otic and olfactory placodes, and in all classes of spinal
neurons.
CC -1- DEVELOPMENTAL STAGE: Onset of expression correlates with the
initiation of axogenesis in 16-36 hour embryos.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily.
CC -1- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 5 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X89204; CAA61490.1; .
CC F01; T30581; T30581.
CC HSPF; P20241; 1CFB.
CC ZFIN; ZDB-GENE-980526-512; nadl1.1.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR003962; FNIII_subd.
```

```
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR Pfam; PF00041; fn3; 5.
DR Pfam; PF00047; Ig; 6.
DR PRINTS; PR00014; FNTYPEIII.
DR SMART; SM00060; FN3; 5.
DR SMART; SM00408; IGG2; 4.
DR PROSITE; PS50835; IG LIKE; 6.
KW Neurogenesis; Cell adhesion; Developmental protein; Glycoprotein;
transmembrane; Repeat; Immunoglobulin domain.
FT DOMAIN 1 1054 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1055 1075 POTENTIAL.
FT DOMAIN 1076 1197 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 58 IG-LIKE C2-TYPE 1.
FT DOMAIN 69 160 IG-LIKE C2-TYPE 2.
FT DOMAIN 165 263 IG-LIKE C2-TYPE 3.
FT DOMAIN 268 355 IG-LIKE C2-TYPE 4.
FT DOMAIN 360 442 IG-LIKE C2-TYPE 5.
FT DOMAIN 451 541 IG-LIKE C2-TYPE 6.
FT DOMAIN 546 638 FIBRONECTIN TYPE-III 1.
FT DOMAIN 645 739 FIBRONECTIN TYPE-III 2.
FT DOMAIN 744 849 FIBRONECTIN TYPE-III 3.
FT DOMAIN 850 948 FIBRONECTIN TYPE-III 4.
FT DOMAIN 952 1029 FIBRONECTIN TYPE-III 5.
FT DISULFID 92 143 BY SIMILARITY.
FT DISULFID 199 247 BY SIMILARITY.
FT DISULFID 289 339 BY SIMILARITY.
FT DISULFID 383 432 BY SIMILARITY.
FT DISULFID 472 525 BY SIMILARITY.
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 708 708 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 959 959 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 968 968 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1002 1002 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1197 AA; 132860 MW; 7CE1505EEFAC7B28 CRC64;

Query Match 14.0%; Score 88.5; DB 1; Length 1197;
Best Local Similarity 25.9%; Pred. No. 0.34; Mismatches 17; Gaps 4;
Matches 29; Conservative 14; Indels 52;

QY 3 VSQPEIRTELGSSAFPCSFNAGSLAIGSVTWFRDDEVVPGKEVNRGTPEFRGLAPL 62
Db 454 VGPQNHLVIRGSDAILHCKYTVDH-NLKSPTVQWKKD-----GHKITAST----- 498
QY 63 ASSRFLDHQALHVRDGHDSIYVCRVEVLGLGVGTGNGTRLVVEKEHP 114
Db 499 -SNDKTHEIGSLKVLVDQVQMDMGISYCEVSTT-LDSDTASGVITVQDKPDP 548

RESULT 8
C79A_HUMAN
ID C79A_HUMAN STANDARD; PRT; 226 AA.
AC P11912;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE B-cell antigen receptor complex associated protein alpha-chain
precursor (Ig-alpha) (MB-1 membrane glycoprotein) (Surface-IgM-
associated protein) (Membrane-bound immunoglobulin associated
protein) (CD79a).
DE CD79A OR IGA OR MB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92166394; PubMed=1538135;  
 RA Ha H.J., Kubagawa H., Burrows P.D.;  
 RT "Molecular cloning and expression pattern of a human gene homologous  
 RT to the murine mb-1 gene.";  
 RL J. Immunol. 148:1526-1531(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92105765; PubMed=1729378;  
 RA Yu L.M., Chang T.W.;  
 RT "Human mb-1 gene: complete cDNA sequence and its expression in B  
 RT cells bearing membrane Ig of various isotypes.";  
 RL J. Immunol. 148:633-637(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93009083; PubMed=1395095;  
 RA Leduc I., Preud'Homme J.L., Cogne M.;  
 RT "Structure and expression of the mb-1 transcript in human lymphoid  
 RT cells.";  
 RL Clin. Exp. Immunol. 90:141-146(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=92347937; PubMed=1639443;  
 RA Flawinkel H., Reth M.;  
 RT "Molecular cloning of the Ig-alpha subunit of the human B-cell  
 RT antigen receptor complex.";  
 RL Immunogenetics 36:266-269(1992).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Tonsil;  
 RX MEDLINE=92289825; PubMed=1534761;  
 RX Mueller B.S., Cooper L., Terhorst C.;  
 RT "Cloning and sequencing of the cDNA encoding the human homologue of  
 RT the murine immunoglobulin-associated protein B29.";  
 RL Eur. J. Immunol. 22:1621-1625(1992).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RA Koyama M., Nakamura T.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP PRELIMINARY SEQUENCE OF 190-226 FROM N.A.  
 RA MEDLINE=89091088; PubMed=2463161;  
 RA Sakaguchi N., Kashiwamura S., Kimoto M., Thalmann P., Melchers F.;  
 RT "B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-  
 RT like structural properties.";  
 RL EMBO J. 7:3457-3464(1988).  
 RN [8]  
 RP SEQUENCE OF 33-52.  
 RX MEDLINE=94239426; PubMed=7514267;  
 RX Vasile S., Coligan J.E., Yoshida M., Seon B.K.;  
 RT "Isolation and chemical characterization of the human B29 and mb-1  
 RT proteins of the B cell antigen receptor complex.";  
 RL Mol. Immunol. 31:419-427(1994).  
 CC -!- FUNCTION: Associated to surface Igm-receptor; may be involved in  
 CC signal transduction.  
 CC -!- SUBUNIT: Heterodimer of alpha and beta chains; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Long;  
 CC IsoId=Fl1912-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=Fl1912-2; Sequence=VSP\_002476;  
 CC -!- TISSUE SPECIFICITY: B-cells.  
 CC -!- PTM: Phosphorylated on tyrosine; upon B-cell activation.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD Guide CD79a entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd79a.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL; U05259; AAA20495.1; -;  
 DR EMBL; M86921; AAA59557.1; -;  
 DR EMBL; X13451; CRA31802.1; ALT\_SEQ.  
 DR EMBL; S75217; AAB20812.1; -;  
 DR EMBL; S46706; AAB23558.1; -;  
 DR EMBL; M80462; AAA59556.1; -;  
 DR EMBL; X83440; CAA58523.1; -;  
 DR EMBL; M74721; AAA60270.1; -;  
 DR PIR; I54539; A46477.  
 DR Genew; HGNC:1698; CD79A.  
 DR MIM; 112205; -;  
 DR GO; GO:0007166; P:cell surface receptor linked signal transdu. .; TAS.  
 DR GO; GO:0006952; P:defense response; TAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR003110; ITAM.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF02189; ITAM; 1.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00077; ITAM; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR Immunoglobulin domain; Transmembrane; Glycoprotein; Signal;  
 KW Phosphorylation; Alternative splicing.  
 FT SIGNAL 1 32  
 FT CHAIN 33 226  
 FT B-CELL ANTIGEN RECEPTOR COMPLEX  
 FT ASSOCIATED PROTEIN ALPHA-CHAIN.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT IG-LIKE C2-TYPE.  
 FT POTENTIAL.  
 FT INTERCHAIN (WITH C-136 IN BETA CHAIN)  
 FT (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT GTLIQNVKSHGGIYVCRVQEGNESYQQSCGYLVRVQ  
 FT -> E (in isoform Short).  
 FT /FtId=VSP\_002476.  
 FT V -> I (IN REF. 4).  
 FT CONFLICT 69 69  
 FT SEQUENCE 226 AA; 25038 MW; 6E5B837409969292 CRC64;  
 Query Match 13.9%; Score 88; DB 1; Length 226;  
 Best Local Similarity 29.9%; Pred. No. 0.061;  
 Matches 38; Conservative 13; Mismatches 50; Indels 26; Gaps 7;  
 QY 1 LWSQ-QPEIRTEGSSAFPCSFNASQGLAISGVTWFRDEVVPGKEVNGT--PEPRG 57  
 DB 33 LMMKVPASLAVSLGSDAHFQCPHNSNN---ANVTWVR--VLHG----NYTWPFPELG 82  
 QY 58 RLAPLASSRFLHDHQAELHIDVRGHDAIYVCRVEVLGLGVGTGNGTRLVVEKEHPQ-- 115  
 DB 83 P-----GEDPNGTLLIQNVKSHGGIYVCRVQEGNESYQQSCGYLVRVQPPRPF 133  
 QY 116 --LGAGT 120  
 DB 134 LDMGEGT 140  
 RESULT 9  
 ID BUTY\_MOUSE STANDARD; PRT; 524 AA.  
 AC Q62556; P97392;  
 DT 01-NOV-1997 (Rel. 35, Created)



RT "Complete sequence determination of the mouse and human CTLA4 gene  
 RT loci, cross-species DNA sequence similarity beyond exon borders.";   
 RL Genomics 60:341-355 (1999).   
 RN [6]   
 RN FUNCTION.   
 RP MEDLINE=91341416; PubMed=1714933;   
 RX Linsley P.S., Brady W., Urnes M., Griensmaire L.S., Danle N.K.,   
 RA Ledbetter J.A.;   
 RT "CTLA-4 is a second receptor for the B cell activation antigen B7.";   
 RL J. Exp. Med. 174:561-569 (1991).   
 RN [7]   
 RN STRUCTURE BY NMR OF 37-165.   
 RP MEDLINE=97372889; PubMed=228944;   
 RX Metzler W.J., Bajjorath J., Fenderson W., Shaw S.Y., Constantine K.L.,   
 RA Naemura J., Leytze G., Peach R.J., Lavoie T.B., Mueller L.,   
 RA Linsley P.S.;   
 RT "Solution structure of human CTLA-4 and delineation of a CD80/CD86   
 RT binding site conserved in CD28.";   
 RL Nat. Struct. Biol. 4:527-531 (1997).   
 RN [8]   
 RN VARIANT ALA-17, AND ASSOCIATION WITH IDDM12.   
 RP MEDLINE=97402209; PubMed=925273;   
 RX Marron M.P., Raffel L.J., Garchon H.-J., Jacob C.O., Serrano-Rios M.,   
 RA Martinez Larrad M.T., Teng W.-P., Park Y., Zhang Z.-X.,   
 RA Goldstein D.R., Tao Y.-W., Beaurain G., Bach J.-F., Huang H.-S.,   
 RA Luo D.-F., Zeidler A., Rottier J.I., Yang M.C.K., Modilevsky T.,   
 RA MacLaren N.K., She J.-X.;   
 RT "Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA4   
 RT polymorphisms in multiple ethnic groups.";   
 RL Hum. Mol. Genet. 6:1275-1282 (1997).   
 RN [9]   
 RN POLYMORPHISM, AND ASSOCIATION WITH COELIAC DISEASE.   
 RP MEDLINE=99205840; PubMed=10189842;   
 RX Djilali-Saiah I., Schmitz J., Harfouch-Hamoud E., Mougenot J.-F.,   
 RA Bach J.-F., Caillaud-Zucman S.;   
 RT "CTLA-4 gene polymorphism is associated with predisposition to coeliac   
 RT disease.";   
 RL Gut 43:187-189 (1998).   
 RN [10]   
 RN VARIANT ALA-17, AND ASSOCIATION WITH TAO.   
 RP MEDLINE=99402177; PubMed=10475192;   
 RX Vaidya B., Imrie H., Perros P., Dickinson J., McCarthy M.I.,   
 RA Kendall-Taylor P., Pearce S.H.S.;   
 RT "Cytotoxic T lymphocyte antigen-4 (CTLA-4) gene polymorphism confers   
 RT susceptibility to thyroid associated orbitopathy.";   
 RL Lancet 354:743-744 (1999).   
 RN [11]   
 RN VARIANT ALA-17, AND ASSOCIATION WITH GRD.   
 RP MEDLINE=20385252; PubMed=10924276;   
 RX Chistyakov D.A., Savost'yanov K.V., Turakulov R.I., Petunina N.A.,   
 RA Trukhina L.V., Kudanova A.V., Balabolkin M.I., Nosikov V.V.;   
 RT "Complex association analysis of Graves disease using a set of   
 RT polymorphic markers.";   
 RL Mol. Genet. Metab. 70:214-218 (2000).   
 RN [12]   
 RN VARIANT ALA-17.   
 RP MEDLINE=20395844; PubMed=10903931;   
 RX Deng Z., Morse J.H., Slager S.L., Cuervo N., Moore K.J., Venetos G.,   
 RA Kalachikov S., Cayanis E., Fischer S.G., Barst R.J., Hodge S.E.,   
 RA Knowles J.A.;   
 RT "Familial primary pulmonary hypertension (gene PPH1) is caused by   
 RT mutations in the bone morphogenetic protein receptor-II gene.";   
 RL Am. J. Hum. Genet. 67:737-744 (2000).   
 CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1   
 CC (CD80) AND B7-2 (CD86).   
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.   
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in   
 CC lymphoid tissues.   
 CC -1- POLYMORPHISM: The variant Ala-17 is associated with an increased   
 CC risk for autoimmune disorders as Graves disease (GRD), type I   
 CC insulin-dependent diabetes mellitus (IDDM12), and thyroid-   
 CC associated orbitopathy (TAO). The variant Thr-17 is associated   
 CC with predisposition to coeliac disease, a gluten sensitive

enteropathy characterized by small bowel mucosal atrophy.   
 -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.   
 -1- DATABASE: NAME=PROW; NOTE=CD guide CD152 entry;   
 WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd152.htm".   
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration   
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -   
 CC the European Bioinformatics Institute. There are no restrictions on its   
 CC use by non-profit institutions as long as its content is in no way   
 CC modified and this statement is not removed. Usage by and for commercial   
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/   
 CC or send an email to license@isb-sib.ch).   
 CC -----

EMBL; AF411058; AAL40932.1; -   
 DR EMBL; AF411058; AAL40932.1; -   
 DR EMBL; M74363; AAA52127.1; -   
 DR EMBL; M37245; AAA52773.1; -   
 DR EMBL; M37245; AAA52773.1; JOINED.   
 DR EMBL; M37244; AAA52773.1; JOINED.   
 DR EMBL; AF142144; AAF02499.1; -   
 DR PIR; S08614; S08614.   
 DR PDB; 1AHL; 15-APR-98.   
 DR PDB; 1H6E; 28-NOV-01.   
 DR PDB; 1I85; 04-APR-01.   
 DR PDB; 1I8L; 04-APR-01.   
 DR Genew; HGNC:2505; CTLA4.   
 DR MIM; 123890; -   
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.   
 DR GO; GO:0006955; P: immune response; TAS.   
 DR InterPro; IPR008096; CTLA4.   
 DR InterPro; IPR007110; IG-like.   
 DR InterPro; IPR003599; IG.   
 DR InterPro; IPR003596; IG V.   
 DR PRINTS; PR01720; CTLANTIGEN4.   
 DR SMART; SM00409; IG; 1.   
 DR SMART; SM00406; IG; 1.   
 DR PROSITE; PS00835; IG LIKE; FALSE NEG.   
 DR Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal;   
 KW 3D-structure; Polymorphism.   
 FT SIGNAL 1 35   
 FT CHAIN 36 223   
 FT DOMAIN 36 161   
 FT TRANSMEM 162 187   
 FT DOMAIN 188 223   
 FT DOMAIN 39 140   
 FT DISULFID 58 129   
 FT DISULFID 85 103   
 FT CARBOHYD 113 113   
 FT VARIANT 17 17   
 FT CONFLICT 147 147   
 FT STRAND 45 47   
 FT TURN 50 51   
 FT STRAND 53 59   
 FT STRAND 70 76   
 FT STRAND 81 88   
 FT TURN 91 92   
 FT STRAND 99 100   
 FT STRAND 104 108   
 FT TURN 109 110   
 FT STRAND 111 117   
 FT TURN 121 123   
 FT STRAND 125 132   
 FT TURN 137 138   
 FT STRAND 142 143   
 FT STRAND 147 150   
 FT SEQUENCE 223 AA; 24656 MW; 6F9466FB2E139A5A CRC64;   
 Query Match 13.8%; Score 87.5; DB 1; Length 223;   
 Best Local Similarity 27.4%; Pred. No. 0.068; 56; Indels 19; Gaps 4;   
 Matches 34; Conservative 15; Mismatches 15;   
 QY 3 VSQPEIRTEGSSAFPCSFNAGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLAPL 62

```
Db 40 VAQPAVAVLASSRGASVCEY-ASPGKATEVRVTVLRQADSQVTEVCAATYMGNELT-- 96
Qy 63 ASRFLHD-----HQLHHRDVRGHDASIVCRVEVL---GLGVGTGNGRLVVE 110
Db 97 ----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMPPPYILGNGTQIYVI 152
Qy 111 KEHP 114
Db 153 DPEP 156

RESULT 11
MOG_BOVIN
ID -MOG BOVIN STANDARD; PRT; 246 AA.
AC P55803;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 63-70.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nussebaum J.-L., Rousset G.,
RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.F.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex."
RN Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994 (1993).
[2]
RP SEQUENCE OF 29-36.
RC TISSUE=Brain;
RX MEDLINE=93382604; PubMed=8371836;
RA Birling M.C., Rousset G., Nussebaum P., Nussebaum J.-L.;
RT "Biochemical and immunohistochemical studies with specific polyclonal
RT antibodies directed against bovine myelin/oligodendrocyte
RT glycoprotein."
RN Neurochem. Res. 18:937-945 (1993).
CC -!- FUNCTION: Minor component of the myelin sheath. May be involved in
CC completion and/or maintenance of the myelin sheath and in cell-
CC cell communication.
CC -!- SUBUNIT: May form homodimers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
CC localized on the surface of myelin and oligodendrocyte cytoplasmic
CC membranes.
CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG
CC family.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
CC with oligodendrocyte-myelin glycoprotein (OMG).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; L21757; -; NOT_ANNOTATED_CDS.
CC DR PIR; A47712; A47712.
CC DR HSRF; Q13740; 1K2C.
CC DR InterPro; IPR007110; Ig-like.
CC DR InterPro; IPR003596; Ig_v.
```

```
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 246 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
FT DOMAIN 29 153 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 154 174 POTENTIAL.
FT DOMAIN 175 209 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 210 230 POTENTIAL.
FT DOMAIN 231 246 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 29 144 IG-LIKE V-TYPE.
FT DISULFID 52 126 POTENTIAL.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 246 AA; 28028 MW; 363C76AB1A33DE41 CRC64;

Query Match 13.8%; Score 87.5; DB 1; Length 246;
Best Local Similarity 34.0%; Pred. No. 0.076;
Matches 32; Conservative 9; Mismatches 34; Indels 19; Gaps 6;

Qy 9 IRTLEGSAFLPCSFNASQGRLAIG-SVTWFRDVEVPGKEV-----RNG-----TPEFR 56
Db 40 IRLVGVDELPC--RISPGKNATGMEVGYRP---PFSRVVHLRYNGKQDEQAPEYR 94
Qy 57 GRLAPLASSRFLHDHQAELHHRDVRGHDASTIYVC 90
Db 95 GRTQLLKEIT--IGEGKVTILRIKRVFSDGEGFTC 126

RESULT 12
MOG_HUMAN
ID -MOG HUMAN STANDARD; PRT; 247 AA.
AC Q16553; Q00713; Q00714; Q00715; Q13054; Q13055; Q14855; Q92891;
AC Q92892; Q92893; Q92894; Q92895; Q93053; Q96K03; Q96K05;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=93310943; PubMed=7790876;
RA Hilton A.A., Slavin A.J., Hilton D.J., Bernard C.C.A.;
RT "Characterization of cDNA and genomic clones encoding human myelin
RT oligodendrocyte glycoprotein."
RN J. Neurochem. 65:309-318 (1995).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95054056; PubMed=7964757;
RA Pham-Dinh D., Allinquant B., Ruberg M., della Gaspera B.,
RA Nussebaum J.-L., Dautigny A.;
RT "Characterization and expression of the cDNA coding for the human
RT myelin/oligodendrocyte glycoprotein."
RN J. Neurochem. 63:2353-2356 (1994).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96015053; PubMed=8530032;
RA Roth M.-P., Malfroy L., Offer C., Sevin J., Enault G., Borot N.,
RA Pontarotti P., Coppin H.;
RT "The human myelin oligodendrocyte glycoprotein (MOG) gene: complete
RT nucleotide sequence and structural characterization."
RN Genomics 28:241-250 (1995).
[4]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=96115584; PubMed=8666381;
RA Pham-Dinh D., della Gaspera B., de Rosbo N.K., Dautigny A.;
RT "Structure of the human myelin/oligodendrocyte glycoprotein gene and
RT multiple alternative spliced isoforms."
```

Genomics 29:345-352 (1995).

[5] SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.

RA Ballenthin P.A., Gardiner M.V., to the EMBL/GenBank/DBSJ databases.

[6] SEQUENCE FROM N.A. (ISOFORMS 1; 5 AND 7).

RA Griffiths C.,

RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.

[7] STRUCTURE BY NMR OF 54-84.

RP MEDLINE=97354172; PubMed=9210466;

RA Albouze-Abo S., Wilson J.C., Bernard C.C.A., von Itzstein M.,

RT "A conformational study of the human and rat encephalitogenic myelin

RT oligodendrocyte glycoprotein peptides 35-55."

RL Eur. J. Biochem. 246:59-70 (1997).

CC -!- FUNCTION: Minor component of the myelin sheath. May be involved in

CC completion and/or maintenance of the myelin sheath and in cell-

CC cell communication.

CC -!- SUBUNIT: May form homo- or heterodimers between the different

CC isoforms.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (isoforms 1 and

CC 5); type I membrane protein (other isoforms) (Potential).

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=9;

CC Comment=Additional isoforms seem to exist;

CC Name=1; Synonyms=Alpha-1;

CC IsoId=Q16653-1; Sequence=Displayed;

CC Name=2; Synonyms=Alpha-2;

CC IsoId=Q16653-2; Sequence=VSP\_002543;

CC Name=3; Synonyms=Alpha-3;

CC IsoId=Q16653-3; Sequence=VSP\_002542;

CC Name=4; Synonyms=Alpha-4;

CC IsoId=Q16653-4; Sequence=VSP\_002539;

CC Name=5; Synonyms=Beta-1;

CC IsoId=Q16653-5; Sequence=VSP\_002545;

CC Name=6; Synonyms=Beta-2;

CC IsoId=Q16653-6; Sequence=VSP\_002543;

CC Name=7; Synonyms=Beta-3;

CC IsoId=Q16653-7; Sequence=VSP\_002542;

CC Name=8; Synonyms=Beta-4;

CC IsoId=Q16653-8; Sequence=VSP\_002544;

CC Name=9;

CC IsoId=Q16653-9; Sequence=VSP\_002540;

CC Note=Not functionally active. May be expressed at low level in

CC the adult;

CC -!- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is

CC localized on the surface of myelin and oligodendrocyte cytoplasmic

CC membranes.

CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG

CC family.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC -!- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)

CC with oligodendrocyte-myelin glycoprotein (OMG).

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; X74511; CAAS2617.1; -

DR EMBL; Z48051; CAAB8109.1; -

DR EMBL; U18840; AAC50361.1; -

DR EMBL; U18843; AAC50362.1; -

DR EMBL; U18798; AAC50875.1; -

DR EMBL; U18799; AAC50877.1; -

DR EMBL; U18800; AAC36870.1; -

DR EMBL; U18801; AAC50878.1; -

DR EMBL; U18803; AAC50879.1; -

DR EMBL; U64564; AAB08088.1; -

DR EMBL; U64565; AAB08089.1; -

DR EMBL; U64566; AAB08090.1; -

DR EMBL; U64567; AAB08091.1; -

DR EMBL; U64568; AAB08092.1; -

DR EMBL; U64569; AAB08093.1; -

DR EMBL; U64570; AAB08094.1; -

DR EMBL; U64571; AAB08095.1; ALT\_SEQ.

DR EMBL; AL050328; CAB89267.1; -

DR EMBL; AL050328; CAB89269.1; -

DR EMBL; AL050328; CAB89270.1; -

DR EMBL; AL050328; CAB89271.1; -

DR Genew; HGNC:7197; MOG.

DR MIN; 159465; -

DR GO; GO:0007417; P:Central nervous system development; TAS.

DR InterPro; IPR007110; Ig-like.

DR Pfam; PF00047; Ig; 1.

DR PROSITE; PS00835; IG\_LIKE; 1.

KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal;

KW Alternative splicing.

FT SIGNAL 1 29 POTENTIAL.

FT CHAIN 30 247 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.

FT DOMAIN 30 154 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 155 175 POTENTIAL.

FT DOMAIN 176 210 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 211 231 POTENTIAL.

FT DOMAIN 232 247 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 32 145 IG-LIKE V-TYPE.

FT DISULFID 53 127 POTENTIAL.

FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 30 145 Missing (in isoform 4).

FT VARSPLIC 198 203 /FTid=VSP\_002539.

FT VARSPLIC 198 203 DPHPLR -> GKPRHV (in isoform 9).

FT VARSPLIC 204 247 /FTid=VSP\_002540.

FT VARSPLIC 198 236 Missing (in isoform 9).

FT VARSPLIC 198 236 /FTid=VSP\_002541.

FT VARSPLIC 198 236 DPHPLRVCWKITLFIIVPLVPLVALIICYNLHRLA

FT VARSPLIC 244 247 -> ESFGVLGPQVKEPKKT (in isoform 3 and

FT VARSPLIC 198 203 isoform 7).

FT VARSPLIC 198 236 /FTid=VSP\_002542.

FT VARSPLIC 198 243 Missing (in isoform 2 and isoform 6).

FT VARSPLIC 198 243 /FTid=VSP\_002543.

FT VARSPLIC 244 247 Missing (in isoform 8).

FT VARSPLIC 171 171 RNP -> LFHLBALSG (in isoform 5, isoform

FT CONFLICT 171 171 6, isoform 7 and isoform 8).

FT SEQUENCE 247 AA; 28179 MW; 847601FE5937AB0C CRC64;

Query Match 13.8%; Score 87.5; DB 1; Length 247;

Best Local Similarity 34.0%; Pred. No. 0.076;

Matches 32; Conservative 9; Mismatches 34; Indels 19; Gaps 6;

QY 9 IRTLEGSSAFPCSFNASQGLAIG-SVTWFRDEVVPGKEV----RNG-----TPEPR 56

DB 41 IRLVGVDEVSLPC--RISPGKNATGMEVGIWP---PFSRVHLYRNGKQDGDQAPYR 95

QY 57 GRPLAPLASSRFLHDHQAELHVRDVRGHDAIYVC 90

DB 96 GRTELLAKDA--IGBKVTLRINRVFSDGGFTC 127

RESULT 13

CTLA4\_PIG STANDARD; PRT; 223 AA.

ID CTLA4\_PIG

AC Q9MYX7;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-

DE associated antigen 4) (CTLA-4).

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OCC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 NCBI\_TaxID=9823;  
 [1]  
 RN  
 RP SEQUENCE FROM N.A.  
 RRT Tachedian M., Chaplin P.J., Scheerlinck J.-P.Y., Tennent J.M.;  
 RRT "Molecular characterization and phylogenetic analysis of porcine  
 RRT cytotoxic T-lymphocyte-associated antigen 4 (CTLA4)".  
 RRT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1  
 CC (CD80) AND B7-2 (CD86) (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collaborator -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF281633; AAF86981.1; -;  
 CC HSP: P16410; IAH1.  
 DR InterPro: IPR008096; CTLA4.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003599; IG.  
 DR PRINTS: PR01720; CTLANTIGEN4.  
 DR SMART: SM00409; IG; 1.  
 DR PROSITE: PS50835; IG LIKE; FALSE NEG.  
 CC Immunoglobulin domain; T-cell, T-membrane; Glycoprotein; Signal.  
 FT SIGNAL 1 35  
 FT CHAIN 36 223  
 FT DOMAIN 38 161  
 FT TRANSMEM 162 187  
 FT DOMAIN 188 223  
 FT DOMAIN 39 140  
 FT DISULFID 58 129  
 FT DISULFID 85 103  
 FT CARBOHYD 113 113  
 FT CARBOHYD 145 145  
 FT SEQUENCE 223 AA; 24430 MW; 2DD865D871D8921 CRC64;  
 SQ  
 Query Match 13.5%; Score 85.5; DB 1; Length 223;  
 Best Local Similarity 28.2%; Pred. No. 0.11;  
 Matches 35; Conservative 14; Mismatches 56; Indels 19; Gaps 4;  
 QY 3 VQSPPEITLEGSSAFLEPCSFNAGSGLRATGVTWTFRDEVPVPGKEVRNGTPEFRGLAPL 62  
 DB :|||:  
 DB 40 VAQPAVTLANSRGVASFVCEY-GSAGKAAEVRVTLFRAGSQMTVECAATYVEDELT- 96  
 QY 63 ASSRFIHD-----HQAELHTRDVRGHDDASTYVCRVEVL---GLGVGTGNGTRLVVE 110  
 DB :|||:  
 DB 97 ----FLLDSTCTGTSTENKVNLTQGLRAVDTSGLYICKVELLYPPPYVTGMGNGTQIVY 152  
 QY :|||:  
 DB :|||:  
 DB 111 KGHF 114  
 DB 153 DPEP 156

NCBI\_TaxID=9985;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=B/J X CHBS:HM;  
RX MEDLINE=95369849; PubMed=7642234;  
RA Isono T., Seto A.;  
RT "Cloning and sequencing of the rabbit gene encoding T-cell  
RL costimulatory molecules.";  
RL Immunogenetics 42:217-220(1995).  
CC CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1  
(CD80) AND B7-2 (CD86) (BY SIMILARITY).  
CC CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D49844; BAA08644.1; --  
CC PIR; I46696; I46696.  
CC HSP; P16410; IAH1.  
DR DR InterPro; IPR008096; CTLA4.  
DR DR InterPro; IPR007110; Ig-like.  
DR DR InterPro; IPR003599; Ig.  
DR DR PRINTS; PRO1720; CTLANTIGEN4.  
DR DR SMART; SM00409; IG; 1.  
DR DR PROSITE; PS08935; IG\_LIKE; 1.  
KW Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 35  
FT CHAIN 36 223  
FT DOMAIN 38 161  
FT TRANSMEM 162 187  
FT DOMAIN 188 223  
FT CHAIN 36 145  
FT DOMAIN 58 129  
FT DISULFID 85 103  
FT CARBOHYD 113 113  
FT CARBOHYD 145 145  
SQ SEQUENCE 223 AA; 24655 MW; 85A9269793B88160 CRC64;  
  
Query Match 13.4%; Score 84.5; DB 1; Length 223;  
Best Local Similarity 27.5%; Pred. No. 0.14;  
Matches 33; Conservative 16; Mismatches 64; Indels 7; Gaps 3;  
  
QY 1 LWSQAPPEIRTEGSSAFLEPCSFNASQGRLAGISVTWFRDEVVPGKVRNGTPEFRGLA 60  
Db 38 LHSQPAVVLASSRGSVAFVCEY-ASSHKATEVRVTVLRQANSQMTVCAMTYTVNELT 96  
QY 61 PLASSR---FLHDHQAEHLHEDVRGHDSALYVCEVEVL---GLGVGTGNGTRLVVEKEHP 114  
Db 97 FIDSDCTGTGSHGNKNVLTIQGSAMDTGLYICKVELMWPPPYVVGNGTQIIVIEPEP 156  
  
RESULT 15  
MYPO HETFP  
ID MYPO HETFP STANDARD; PRT; 246 AA.  
AC 220936;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral  
DE protein) (MPP).  
OS Heterodontus francisci (Horn shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Heterodontoidea; Heterodontiformes;  
OC Heterodontidae; Heterodontus.  
OX NCBI\_TaxID=7792;  
[1]  
RP SEQUENCE FROM N.A.



```

RC TISSUE=Brain;
RX MEDLINE=90040744; PubMed=2478717;
RA Saavedra R.A., Fors L., Aebersold R.H., Arden B., Horvath S.,
RA Sanders J., Hood L.;
RT "the myelin proteins of the shark brain are similar to the myelin
RT proteins of the mammalian peripheral nervous system.";
RL J. Mol. Evol. 29:149-156(1989).
CC CC -1- FUNCTION: Creation of an extracellular membrane face which guides
CC the wrapping process and ultimately compacts adjacent lamellae.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Found only in peripheral nervous system
CC Schwann cells.
CC CC -1- FTM: N-GLYCAN IS SULFATED (BY SIMILARITY).
CC CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC CC -1- SIMILARITY: Belongs to the myelin P0 protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15714; CAB37865.1; --
DR PIR; A32999; A32999.
DR HSP; P06907; INEU.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_V.
DR InterPro; IPR000920; Myelin_P0.
DR PRINTS; PR00213; MYELINP0.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00568; MYELIN_P0; 1.
KW Myelin; Structural protein; Glycoprotein; Transmembrane;
KW Phosphorylation; Immunoglobulin domain; Signal.
FT SIGNAL 1 27
FT CHAIN 28 246 MYELIN P0 PROTEIN.
FT DOMAIN 28 150 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 151 178 POTENTIAL.
FT DOMAIN 179 246 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 143 IG-LIKE V-TYPE.
FT DISULFID 48 125 POTENTIAL.
FT CARBOHYD 120 120 N-LINKED (GLCNAC. .) (COMPLEX) (BY
FT SIMILARITY).
SQ SEQUENCE 246 AA; 27335 MW; A776A9ED9D430FA0 CRC64;
Query Match 13.1%; Score 83; DB 1; Length 246;
Best Local Similarity 25.0%; Pred No. 0.22;
Matches 32; Conservative 23; Mismatches 59; Indels 14; Gaps 4;
Qy 3 VSQPEIRTLGGSAFLPCSFNASQGRLAIGSVTW-FRD-----EVPFGKEVRNGTP----- 53
Db 30 VSTHNHLKTVGSDVTLVYCGFWSEVYSDLTTLTSWRFRPNDSRDIISIFHYGNGVPIEK 89
Qy 54 --EERGLAPLASSRFLHDHQAELHNRDVRGHDSIVYCVREVGLGVTGNGTRLVVEK 111
Db 90 WQPRGRVETWGD---ISKHGSIVIRNLDYIDNGTFCDDVKNPDDVWVSSDVHLTYD 146
Qy 112 EHPQLGAG 119
Db 147 KIPPVGAG 154

```

Search completed: February 26, 2004, 12:10:47  
Job time : 11.1857 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 16.1273 Seconds  
(without alignments)  
613.452 Million cell updates/sec

Title: US-10-036-444-2

Perfect score: 1020

Sequence: 1 MAWMLLLILIMVHPSGALW.....GTHCHSDGPRGVPEPCP 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	112	11.0	223	1	CTL4_HUMAN
2	111.5	10.9	226	1	CTL4_HUMAN
3	108.5	10.6	223	1	CTL4_MOUSE
4	106	10.4	132	1	TVB2_MOUSE
5	103.5	10.1	223	1	CTL4_PIG
6	103.5	10.1	223	1	CTL4_RABIT
7	100.5	9.9	1694	1	SN_MOUSE
8	99	9.7	467	1	SL15_MOUSE
9	98.5	9.7	526	1	BUTY_HUMAN
10	98	9.6	245	1	MOG_RAT
11	96.5	9.5	246	1	MOG_MOUSE
12	96	9.4	131	1	LV6E_HUMAN
13	95	9.3	524	1	BUTY_MOUSE
14	93.5	9.2	247	1	MOG_HUMAN
15	93	9.1	246	1	MPPO_HETFR
16	92.5	9.1	403	1	CD33_MOUSE
17	92	9.0	215	1	CTB2_HUMAN
18	90.5	8.9	246	1	MOG_BOVIN
19	89	8.7	142	1	VPB2_MOUSE
20	89	8.7	526	1	BUTY_BOVIN
21	89	8.7	628	1	LU_HUMAN
22	89	8.7	1709	1	SN_HUMAN
23	88.5	8.7	1197	1	CAM1_BRARE
24	88	8.6	463	1	SL19_HUMAN
25	88	8.6	579	1	ASO_CUCNA
26	85.5	8.4	3707	1	PCBM_MOUSE
27	85	8.3	2132	1	PCCA_MOUSE
28	84	8.2	133	1	TVB2_HUMAN
29	84	8.2	2415	1	PCCA_HUMAN
30	83	8.1	134	1	TVB7_MOUSE
31	83	8.1	270	1	BASI_RABIT
32	83	8.1	587	1	ASO_CUCA
33	83	8.1	2333	1	PCCA_CANFA

34	82.5	8.1	597	1	SILL_PANTR
35	82.5	8.1	646	1	MU18_HUMAN
36	82	8.0	365	1	CKAR_HUMAN
37	82	8.0	551	1	SIL5_HUMAN
38	82	8.0	626	1	MAG_HUMAN
39	82	8.0	2124	1	PGC3_RAT
40	81.5	8.0	223	1	C79A_BOVIN
41	81	7.9	552	1	ASO_CUCPM
42	80.5	7.9	365	1	CKAR_MOUSE
43	80	7.8	142	1	VPB1_MOUSE
44	79.5	7.8	130	1	LV1G_HUMAN
45	79.5	7.8	499	1	SIL6_HUMAN

ALIGNMENTS

RESULT 1

CTL4\_HUMAN  
ID CTL4\_HUMAN STANDARD; PRT; 223 AA.  
AC P16410; Q96F43; Q9UK99;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-associated antigen 4) (CTLA-4) (CD152 antigen).  
DE CTLA4 OR CD152.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21601151; PubMed=11735222;  
RA Ling V., Wu P.W., Finnerty H.F., Agostino M.J., Graham J.R., Chen S., Jussiff J., Fisk G.J., Miller C.P., Collins M.;  
RT "Assembly and annotation of human chromosome 2q33 sequence containing the CD28, CTLA4, and ICOS gene cluster: analysis by computational, comparative, and microarray approaches."; Genomics 78:155-168(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX Wu P.W., Ling V.;  
RT "Full length sequence of hCTLA4 cDNA."; Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.  
RN [3]  
RP SEQUENCE OF 1-37 FROM N.A.  
RX MEDLINE=9318145; PubMed=1713603;  
RA Harper K., Balzano C., Novier E., Mattei M.-G., Luciani M.F., Golstein P.;  
RT "CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mouse and human as to sequence, message expression, gene structure, and chromosomal location."; J. Immunol. 147:1037-1044(1991).  
RN [4]  
RP SEQUENCE OF 38-223 FROM N.A.  
RX TISSUE-Lymphocytes;  
RA MEDLINE=89120925; PubMed=3220103;  
RA Dariaevach P., Mattei M.-G., Golstein P., Lefranc M.-P.;  
RT "Human Ig superfamily CTLA-4 gene: chromosomal localization and identity of protein sequence between murine and human CTLA-4 cytoplasmic domains."; Eur. J. Immunol. 18:1901-1905(1988).  
RN [5]  
RP SEQUENCE OF 140-223 FROM N.A., AND TISSUE SPECIFICITY.  
RX MEDLINE=99425274; PubMed=10493833;  
RA Ling V., Wu P.W., Finnerty H.F., Sharpe A.H., Gray G.S., Collins M.;  
RT "Complete sequence determination of the mouse and human CTLA4 gene loci: cross-species DNA sequence similarity beyond exon borders."; Genomics 60:341-355(1999).  
RN [6]  
RP FUNCTION.  
RX MEDLINE=91341416; PubMed=1714933;

RA Linsley P.S., Brady W., Urnes M., Griesmaire L.S., Damle N.K.,  
 RA Ledbetter J.A.;  
 RA "CTLA-4 is a second receptor for the B cell activation antigen B7.",  
 RL J. Exp. Med. 174:561-569 (1991).  
 [7]  
 RP STRUCTURE BY NMR OF 37-165.  
 RA MEDLINE=97372889; PubMed=9228944;  
 RA Metzler W.J., Bajorath J., Henderson M., Shaw S.Y., Constantine K.L.,  
 RA Nemura J., Leytze G., Peach R.J., Lavoie T.B., Mueller L.,  
 RA Linsley P.S.;  
 RA "Solution structure of human CTLA-4 and delineation of a CD80/CD86  
 RT binding site conserved in CD28.",  
 RL Nat. Struct. Biol. 4:527-531 (1997).  
 [8]  
 RP VARIANT ALA-17, AND ASSOCIATION WITH IDDM12.  
 RX MEDLINE=97402209; PubMed=9259273;  
 RA Marron M.P., Raffel L.J., Garchon H.-J., Jacob C.O., Serrano-Rios M.,  
 RA Martinez Larrad M.F., Teng W.-P., Park Y., Zhang Z.-X.,  
 RA Goldstein D.R., Tao Y.-W., Beaurain G., Bach J.-F., Huang H.-S.,  
 RA Luo D.-F., Zeidler A., Rotter J.I., Yang M.C.K., Modilevsky T.,  
 RA MacLaren N.K., She J.-X.;  
 RA "Insulin-dependent diabetes mellitus (IDDM) is associated with CTLA4  
 RT polymorphisms in multiple ethnic groups.",  
 RL Hum. Mol. Genet. 6:1275-1282 (1997).  
 [9]  
 RP POLYMORPHISM, AND ASSOCIATION WITH COELIAC DISEASE.  
 RX MEDLINE=99205840; PubMed=10189842;  
 RA Djilali-Saiah I., Schmitz J., Harfouch-Hamoud E., Mougnot J.-F.,  
 RA Bach J.-F., Caillaud-Zucman S.;  
 RA "CTLA-4 gene polymorphism is associated with predisposition to coeliac  
 RT disease.",  
 RL Lancet 354:743-744 (1999).  
 [10]  
 RP VARIANT ALA-17, AND ASSOCIATION WITH TAO.  
 RX MEDLINE=99402177; PubMed=10475192;  
 RA Valdiva B., Imrie H., Perros P., Dickinson J., McCarthy M.I.,  
 RA Kendall-Taylor P., Pearce S.H.S.;  
 RA "Cytotoxic T lymphocyte antigen-4 (CTLA-4) gene polymorphism confers  
 RT susceptibility to thyroid associated orbitopathy.",  
 RL Lancet 354:743-744 (1999).  
 [11]  
 RP VARIANT ALA-17, AND ASSOCIATION WITH GRD.  
 RX MEDLINE=20385252; PubMed=10924276;  
 RA Chistyakov D.A., Savost'yanov K.V., Turakulov R.I., Petunina N.A.,  
 RA Trukhina L.V., Kudanova A.V., Balabolkin M.I., Nosikov V.V.;  
 RA "Complex association analysis of Graves disease using a set of  
 RT polymorphic markers.",  
 RL Mol. Genet. Metab. 70:214-218 (2000).  
 [12]  
 RP VARIANT ALA-17.  
 RX MEDLINE=20395844; PubMed=10903931;  
 RA Deng Z., Morse J.H., Slager S.L., Cuervo N., Moore K.J., Venetos G.,  
 RA Kalachikov S., Cayanis E., Fischer S.G., Barst R.J., Hodge S.E.,  
 RA Knowles J.A.;  
 RA "Familial primary pulmonary hypertension (gene PPH1) is caused by  
 RT mutations in the bone morphogenetic protein receptor-II gene.",  
 RL Am. J. Hum. Genet. 67:737-744 (2000).  
 CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1  
 CC (CD80) AND B7-2 (CD86).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in  
 CC lymphoid tissues.  
 CC -1- POLYMORPHISM: The variant Ala-17 is associated with an increased  
 CC risk for autoimmune disorders as Graves disease (GRD), type I  
 CC insulin-dependent diabetes mellitus (IDDM12), and thyroid-  
 CC associated orbitopathy (TAO). The variant Thr-17 is associated  
 CC with predisposition to coeliac disease, a gluten sensitive  
 CC enteropathy characterized by small bowel mucosal atrophy.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD152 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd152.htm".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL; AF411058; AAL40932.1; -;  
 CC EMBL; AF414120; AAL07473.1; -;  
 CC EMBL; M74363; AAA52127.1; -;  
 CC EMBL; M37245; AAA52773.1; -;  
 CC EMBL; M37243; AAA52773.1; JOINED.  
 CC EMBL; M37244; AAA52773.1; JOINED.  
 CC EMBL; AF142144; AAF02499.1; -;  
 CC PIR; S08614; S08614.  
 CC PDB; 1AHL1; 15-APR-98.  
 CC PDB; 1H6E; 28-NOV-01.  
 CC PDB; 1I85; 04-APR-01.  
 CC PDB; 1I8L; 04-APR-01.  
 CC Genew; HGNC:2505; CTLA4.  
 CC MIM; 123890; -;  
 CC GO; GO:0005887; C: integral to plasma membrane; TAS.  
 CC GO; GO:0006955; P: immune response; TAS.  
 CC InterPro; IPR008096; CTLA4.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003599; Ig.  
 CC InterPro; IPR003596; Ig v.  
 CC PRINTS; PR01720; CTLANTIGEN4.  
 CC SMART; SM00409; IG; 1.  
 CC SMART; SM00406; IGv; 1.  
 CC PROSITE; PS50835; IG LIKE; FALSE NEG.  
 CC Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal;  
 KW 3D-structure; Polymorphism.  
 FT SIGNAL 1 35  
 FT CHAIN 36 223  
 FT DOMAIN 36 161  
 FT TRANSMEM 162 187  
 FT DOMAIN 188 223  
 FT DOMAIN 39 140  
 FT DISULFID 58 129  
 FT DISULFID 85 103  
 FT CARBOHYD 113 113  
 FT VARIANT 17 17  
 FT CONFLICT 147 147  
 FT STRAND 45 47  
 FT TURN 50 51  
 FT STRAND 53 59  
 FT STRAND 70 76  
 FT STRAND 81 88  
 FT TURN 91 92  
 FT TURN 99 100  
 FT STRAND 104 108  
 FT TURN 109 110  
 FT STRAND 111 117  
 FT TURN 121 123  
 FT STRAND 125 132  
 FT TURN 137 138  
 FT STRAND 142 143  
 FT STRAND 147 150  
 FT SEQUENCE 223 AA; 24556 MW; 6F9466F2E139A5A CRC64;  
 Query Match 11.0%; Score 112; DB 1; Length 223;  
 Best Local Similarity 28.0%; Pred. No. 0.0038;  
 Matches 47; Conservative 22; Mismatches 75; Indels 24; Gaps 6;  
 QY 6 LLLILMHPGSC-ALWYSOPPEIRLEBSSAFPCSFNASQGLAIGSVTFWFRDEVVPGK 64  
 DB 24 LLFFLLFIPVFCAMHQAHPAVILASSRGIAFVCEY-ASPKATEVTVVLRQDSQVT 82  
 QY 65 EYVNGTPEFRGRGLAPLASSRFLPHD-----HQAELHTRDVRGHDASIYCRVYL-- 113  
 DB 83 EVCAATYMGNELT-----FLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVLMYP 136

QY 114 -GLGVCTGNGTRLVWEKEHPQLGAGTVLLRA-----GFVAVSFLSVAV 156  
 Db 137 PPYLIGNGTQIYVIDPEPCDSEFLWILAVSGLFFYSFLTAV 184

RESULT 2  
 C79A\_HUMAN STANDARD; PRT; 226 AA.  
 ID C79A\_HUMAN PRT; 226 AA.  
 AC P11912;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE B-cell antigen receptor complex associated protein alpha-chain  
 DE precursor (Ig-alpha) (MB-1 membrane glycoprotein) (Surface-IgM-  
 DE associated protein) (Membrane-bound immunoglobulin associated  
 DE protein) (CD79a).  
 GN CD79A OR IGA OR MB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9216394; PubMed=1538135;  
 RA Ha H.J.; Kubagawa H.; Burrows P.D.;  
 RT "Molecular cloning and expression pattern of a human gene homologous  
 RT to the murine mb-1 gene.";  
 RL J. Immunol. 148:1526-1531(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92105765; PubMed=1729378;  
 RA Yu L.M.; Chang T.W.;  
 RT "Human mb-1 gene: complete cDNA sequence and its expression in B  
 RT cells bearing membrane Ig of various isotypes.";  
 RL J. Immunol. 148:633-637(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93009083; PubMed=1395095;  
 RA Leduc I.; Preud'Homme J.L.; Cogne M.;  
 RT "Structure and expression of the mb-1 transcript in human lymphoid  
 RT cells.";  
 RL Clin. Exp. Immunol. 90:141-146(1992).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92347937; PubMed=1639443;  
 RA Flaswinkel H.; Reth M.;  
 RT "Molecular cloning of the Ig-alpha subunit of the human B-cell  
 RT antigen receptor complex.";  
 RL Immunogenetics 36:266-269(1992).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92289825; PubMed=1534761;  
 RA Mueller B.S.; Cooper L.; Terhorst C.;  
 RT "Cloning and sequencing of the cDNA encoding the human homologue of  
 RT the murine immunoglobulin-associated protein B29.";  
 RL Eur. J. Immunol. 22:1621-1625(1992).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RA Koyama M.; Nakamura T.;  
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP PRELIMINARY SEQUENCE OF 190-226 FROM N.A.  
 RX MEDLINE=89091088; PubMed=2463161;  
 RA Sakaguchi N.; Kashiwamura S.; Kimoto M.; Thalmann P.; Melchers F.;  
 RT "B lymphocyte lineage-restricted expression of mb-1, a gene with CD3-  
 RT like structural properties.";  
 RL EMBO J. 7:3457-3464(1988).  
 RN [8]  
 RP SEQUENCE OF 33-52.  
 RX MEDLINE=94239426; PubMed=7514267;  
 RA Vasile S.; Coligan J.E.; Yoshida M.; Seon B.K.;

"Isolation and chemical characterization of the human B29 and mb-1  
 proteins of the B cell antigen receptor complex.";  
 Mol. Immunol. 31:419-427(1994).  
 CC -|- FUNCTION: Associated to surface Igm-receptor; may be involved in  
 CC signal transduction.  
 CC -|- SUBUNIT: Heterodimer of alpha and beta chains; disulfide-linked.  
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -|- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=Long;  
 CC IsoId=P11912-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=P11912-2; Sequence=VSP\_002476;  
 CC -|- TISSUE SPECIFICITY: B-cells.  
 CC -|- PTM: Phosphorylated on tyrosine; upon B-cell activation.  
 CC -|- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -|- DATABASE: NAME=PROV; NOTE=CD guide CD79a entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd79a.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; U05259; AAA20495.1; -;  
 CC EMBL; M86921; AAA59557.1; -;  
 CC EMBL; X13451; CAA31802.1; ALT\_SEQ.  
 CC EMBL; S75217; AAB20812.1; -;  
 CC EMBL; M80462; AAB23558.1; -;  
 CC EMBL; M80462; AAB59556.1; -;  
 CC EMBL; X83540; CAA58523.1; -;  
 CC EMBL; M74721; AAA60270.1; -;  
 CC PIR; I54539; A46477.  
 CC Gene; HGNC:1698; CD79A.  
 CC MIM; 112205; -;  
 CC GO; GO:0007166; P:cell surface receptor linked signal transdu. .; TAS.  
 CC GO; GO:0006952; P:defense response; TAS.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003599; IG.  
 CC InterPro; IPR003110; ITAM.  
 CC Pfam; PF00047; IG; 1.  
 CC Pfam; PF02189; ITAM; 1.  
 CC SMART; SM00077; ITAM; 1.  
 CC SMART; SM00077; ITAM; 1.  
 CC PROSITE; PS50635; IG\_LIKE; 1.  
 CC Immunoglobulin domain; Transmembrane; Glycoprotein; Signal;  
 KW Phosphorylation; Alternative splicing.  
 FT SIGNAL 1 32  
 FT CHAIN 33 226  
 FT B-CELL ANTIGEN RECEPTOR COMPLEX  
 FT ASSOCIATED PROTEIN ALPHA-CHAIN.  
 FT EXTRACELLULAR (POTENTIAL).  
 FT POTENTIAL.  
 FT CYTOPLASMIC (POTENTIAL).  
 FT IG-LIKE C2-TYPE.  
 FT POTENTIAL.  
 FT INTERCHAIN (WITH C-136 IN BETA CHAIN)  
 FT (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT GTLIQNVKSHGIVYVCRVQEGNESYQQSCGYLRVQ  
 FT -> E (in isoform Short).  
 FT /FTid=VSP\_002476.  
 FT V -> I (IN REF. 4).  
 FT CONFLICT 69 69  
 FT SEQUENCE 226 AA; 25038 MW; 6B5B83740969292 CRC64;  
 SQ

Query Match 10.9%; Score 111.5; DB 1; Length 226;

Best Local Similarity 28.0%; Pred. No. 0.0043;  
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILLIMVHPSGALWVQ-PEIRITLGGSAFLPCSFNASQRLAIGSVTWFRDEVVP 62  
DB 18 LFLSLAYLGGQALMHWKVPASLMVSLGEDAHFQCPHNSNN-----ANVTWWR--VLH 71  
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDHQAELHIRDVGRHDASIVYCVREVLGLGVGTG 120  
DB 72 G-----NYTWPEFLGP-----GDPNGTLLIQNVKSHGIVYCVQEGNESYQOS 118  
QY 121 NGRLVVEKHPQ-----LGATV-LLRAGFYAVSFLSVAGSTVYVQ 163  
DB 119 CGTYLRVRQPPRPFDFMGEGTKNRIITAEIGIILFCVAVPGTLLIFR 166

RESULT 3  
CTLA4 MOUSE  
ID CTLA4 MOUSE STANDARD; PRT; 223 AA.  
AC P09793; Q9QZ27;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Cytotoxic T-lymphocyte protein 4 precursor (Cytotoxic T-lymphocyte-associated antigen 4) (CTLA-4).  
GN CTLA4 OR CD152.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87258259; PubMed=349540;  
RA Brunet J.-F., Denizot F., Luciano M.-F., Roux-Dosseto M., Suzan M., Mattei M.-G., Golstein P.;  
RT "A new member of the immunoglobulin superfamily -- CTLA-4.";  
RL Nature 328:267-270(1987).  
RN [2]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RX STRAIN=129/SvJ;  
RX MEDLINE=99425274; PubMed=10493833;  
RT Ling V., Wu F.W., Finnerty H.F., Sharpe A.H., Gray G.S., Collins M.;  
RT "Complete sequence determination of the mouse and human CTLA4 gene loci: cross-species DNA sequence similarity beyond exon borders.";  
RL Genomics 60:341-355(1999).  
RN [3]  
RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=91318145; PubMed=1713603;  
RA Harper K., Balzano C., Rouvier E., Mattei M.-G., Luciano M.F., Golstein P.;  
RT "CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mouse and human as to sequence, message expression, gene structure, and chromosomal location.";  
RL J. Immunol. 147:1037-1044(1991).  
CC -1- FUNCTION: POSSIBLY INVOLVED IN T-CELL ACTIVATION. BINDS TO B7-1 (CD80) AND B7-2 (CD86).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: Widely expressed with highest levels in lymphoid tissues.  
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DB EMBL; X05719; CAA29191.1; --  
DB EMBL; AF142145; AAF01489.1; --  
DB EMBL; M74362; AAA37489.1; --  
PIR; A29063; A29063.

DR HSP; P16410; 1RH1.  
DR MGD; MGI:88556; Ctl4.  
DR InterPro; IPR008096; Ctl4.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003599; Ig.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR01720; CTLANTIGEN4.  
DR SMART; SM00409; Ig; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
KW Immunoglobulin domain; T-cell; Transmembrane; Glycoprotein; Signal.  
FT SIGNAL 1 35  
FT CHAIN 36 223  
FT DOMAIN 36 161  
FT TRANSMEM 162 187  
FT DOMAIN 188 223  
FT DOMAIN 36 145  
FT DISULFID 58 129  
FT DISULFID 85 103  
FT CARBOHYD 108 108  
FT CARBOHYD 113 113  
FT CARBOHYD 145 145  
FT CONFLICT 182 182  
SQ SEQUENCE 223 AA; 24993 MW; 5318FAAF416F4685 CRC64;  
Query Match 10.6%; Score 108.5; DB 1; Length 223;  
Best Local Similarity 26.9%; Pred. No. 0.0079;  
Matches 45; Conservative 26; Mismatches 79; Indels 17; Gaps 6;

QY 3 WMLLILIMVHPSGALWVQ-PEIRITLGGSAFLPCSFNASQRLAIGSVTWFRD--- 58  
DB 22 FVALLTLFLFVVFSEAIQVTSVVLASSHGVSFPCPSYSPSHNTDEV-RVTVLRQTNDQ 80  
QY 59 --EVVPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVGRHDASIVYCVREVL--- 113  
DB 81 MTEVCATFTTEKNTVGELD--YPCSGTTF-NESRVNLTIOGLRAVDGLYCKVELMYP 137  
QY 114 GLGVGTGNGRLVVEKHPQLGAGTVL-----LLRAGFYAVSFLSVAV 156  
DB 138 PFVGMGNGTQIYVIDPEPCDSDFLWILWAVSLGLFFYSFLVSAV 184

RESULT 4  
TV2A2 MOUSE  
ID TV2A2 MOUSE STANDARD; PRT; 132 AA.  
AC P01739;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE T-cell receptor alpha chain V region 2B4 precursor.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (CLONE TT11).  
RX MEDLINE=85036634; PubMed=6548551;  
RA Chien Y., Becker D.M., Lindsten T., Okamura M., Cohen D.I., Davis M.M.;  
RT "A third type of murine T-cell receptor gene.";  
RL Nature 312:31-35(1984).  
PIR; A02015; RWSAV.  
DR HSP; P01607; 1REI.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003596; Ig.V.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IG; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
KW T-cell; Receptor; Immunoglobulin domain; Glycoprotein; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 132  
FT DOMAIN 21 113  
FT DOMAIN 114 117  
FT DOMAIN 118 132  
FT T-CELL RECEPTOR ALPHA CHAIN V REGION 2B4.  
FT V SEGMENT.  
FT D SEGMENT.  
FT J SEGMENT.



```

FT DOMAIN 36 145 IG-LIKE V-TYPE.
FT DISULFID 58 129 BY SIMILARITY.
FT DISULFID 85 103 BY SIMILARITY.
FT CARBOHYD 113 113 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 223 AA; 24655 MW; 85A9269793B88160 CRC64;

Query Match 10.1%; Score 103.5; DB 1; Length 223;
Best Local Similarity 28.0%; Pred. No. 0.022;
Matches 45; Conservative 22; Mismatches 83; Indels 11; Gaps 4;

QY 6 LLILMVHPSGSCALVWSOPFIRLESGSFLSCSNASGRLAGSVTWFRDEVVPGKE 65
Db 25 LFSLSFLVFSKALHVSQPAVLASSRGVASFVCEY-ASSHKATREVTVTLRQANSQMT 83
QY 66 VRNGTPEFRGRGLAPLASSR---FLRDHQAEHLHVDVRGHDSIYVCRVEVL---GLGVGT 119
Db 84 VCAMTYTVENELTFIDSDTGTGISHGNKVNLTIGLSAMDTGLYCKVELMYPFVYVGM 143
QY 120 GNGTRLVVEKHPOLGAGTVLLRA----GFYAVSFUSVAV 156
Db 144 GNGTQIVIEPEPCDFSDFLWLILAISGLFFYSPLITAV 184

RESULT 7
SN_MOUSE STANDARD; PRT; 1694 AA.
AC Q62230; Q55216; Q62228; Q62229;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sialoadhesin precursor (Sialic acid binding Ig-like lectin-1) (Siglec-1) (Sheep erythrocyte receptor) (SER).
GN SN OR SA
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Macrophage;
RX MEDLINE=95009950; PubMed=7925291;
RA Crocker P.R., Mucklow S., Boukson V., McWilliam A., Willis A.C.,
RA Gordon S., Milon G., Keim S., Bradfield P.;
RT "Sialoadhesin, a macrophage sialic acid binding receptor for
RT haemopoietic cells with 17 immunoglobulin-like domains.";
RL EMO J. 13:4490-4503(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98051930; PubMed=9383289;
RA Mucklow S., Gordon S., Crocker P.R.;
RT "Characterization of the mouse sialoadhesin gene, Sn.";
RL Mamm. Genome 8:934-937(1997).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=C57BL/6; TISSUE=Spleen;
RX MEDLINE=91266893; PubMed=2050106;
RA Crocker P.R.;
RT "Purification and properties of sialoadhesin, a sialic acid-binding
RT receptor of murine tissue macrophages.";
RL EMO J. 10:1661-1669(1991).
RN [4]
RP SIALIC ACID BINDING.
RX MEDLINE=95179521; PubMed=7533044;
RA Keim S., Pelz A., Schauer R., Filbin M.T., Tang S., de Bellard M.E.,
RA Schnaar R.L., Mahoney J.A., Hartnell A., Bradfield P., Crocker P.R.;
RT "Sialoadhesin, myelin-associated glycoprotein and CD22 define a new
RT family of sialic acid-dependent adhesion molecules of the
RT immunoglobulin superfamily.";
RL Curr. Biol. 4:965-972(1994).
RN [5]
RP BINDING TO SPN.
RX MEDLINE=21136329; PubMed=11238599;

```

```

RA van den Berg T.K., Nath D., Ziltener H.J., Vestweber D., Fukuda M.,
RA van Die I., Crocker P.R.;
RT "CD43 functions as a T cell counterreceptor for the macrophage
RT adhesion receptor sialoadhesin (Siglec-1).";
RL J. Immunol. 166:3637-3640(2001).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS) OF 20-138.
RX MEDLINE=98323385; PubMed=9660955;
RA May A.P., Robinson R.C., Vinson M., Crocker P.R., Jones E.Y.;
RT "Crystal structure of the N-terminal domain of sialoadhesin in
RT complex with 3' sialyllactose at 1.85 A resolution.";
RL Mol. Cell 1:719-728(1998).
RN [7]
RP STRUCTURE OF 20-138 BY NMR, AND MUTAGENESIS OF TRP-21 AND ARG-116.
RX MEDLINE=99321481; PubMed=10393093;
RA Crocker P.R., Vinson M., Kelm S., Drickamer K.;
RT "Molecular analysis of sialoside binding to sialoadhesin by NMR and
RT site-directed mutagenesis";
RL Biochem. J. 341:355-361(1999).
CC -!- FUNCTION: Macrophage-restricted adhesion molecule that mediates
CC sialic-acid dependent binding to lymphocytes, including
CC granulocytes, monocytes, natural killer cells, B-cells and CD8 T-
CC cells (By similarity). Preferentially binds to alpha2,3-linked
CC sialic acid. Binds to SPN/CD43 on T-cells. May play a role in
CC hemopoiesis.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and
CC soluble (isoforms 2 and 3).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q62230-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q62230-2; Sequence=VSP_002573, VSP_002574;
CC Name=3;
CC IsoId=Q62230-3; Sequence=VSP_002575, VSP_002576;
CC -!- TISSUE SPECIFICITY: Expressed by macrophages in various tissues.
CC Highest expression in spleen and lymph node with lower amounts in
CC lung, liver, bone marrow, heart and skin. No expression in thymus,
CC kidney, brain or small intestine.
CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. SIGLEC
CC (sialic acid binding Ig-like lectin) family.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -!- SIMILARITY: Contains 16 immunoglobulin-like C2-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z36293; CAA85290.1; -
CC EMBL; Z36233; CAA85268.1; -
CC EMBL; Z36234; CAA85269.1; -
CC EMBL; U92842; AAB95641.1; -
CC EMBL; U92833; AAB95641.1; JOINED.
CC EMBL; U92834; AAB95641.1; JOINED.
CC EMBL; U92836; AAB95641.1; JOINED.
CC EMBL; U92837; AAB95641.1; JOINED.
CC EMBL; U92838; AAB95641.1; JOINED.
CC EMBL; U92839; AAB95641.1; JOINED.
CC EMBL; U92840; AAB95641.1; JOINED.
CC EMBL; U92841; AAB95641.1; JOINED.
CC FIR; S50065; S50065.
CC PDB; 1QFO; 16-APR-99.
CC PDB; 1QFP; 16-APR-99.
CC MGD; MGI:99668; Sn.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC GO; GO:0005529; F:sugar binding; ISS.
CC GO; GO:0016337; P:cell-cell adhesion; ISS.
CC GO; GO:0007160; P:cell-matrix adhesion; ISS.
CC GO; GO:0006954; P:inflammatory response; ISS.

```





CC	to as the immunoreceptor tyrosine-based inhibitor motif (ITIM).
CC	This motif is involved in downmodulation of cellular responses.
CC	The phosphorylated ITIM motif binds to the SH2 domain of
CC	PTPN6/SHP-1 and PTPN11/SHP-2.
CC	-!- PTM: Phosphorylated on tyrosine residues.
CC	-!- SIMILARITY: Belongs to the immunoglobulin superfamily, SIGLEC
CC	(sialic acid binding Ig-like lectin) family.
CC	-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC	-!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; AF317298; AAG38598.1; -.
DR	EMBL; AF329269; AAK49917.1; -.
DR	MED; MGI:1932475; Siglec5.
DR	GO; GO:0004872; F:receptor activity; IDA.
DR	GO; GO:0005057; F:receptor signaling protein activity; IPI.
DR	GO; GO:0007155; P:cell adhesion; IDA.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003598; Ig_c2.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00047; Ig; 2.
DR	SMART; SM00408; IGC2; 1.
DR	PROSITE; PS00835; IG_LIKE; 3.
DR	PROSITE; PS00290; IG_MHC; 1.
KW	Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
KW	Immunoglobulin domain; Repeat; Phosphorylation.
FT	SIGNAL 1 18
FT	CHAIN 19 467
FT	DOMAIN 19 353
FT	TRANSMEM 354 374
FT	DOMAIN 375 467
FT	DOMAIN 21 141
FT	DOMAIN 152 239
FT	DOMAIN 242 339
FT	DOMAIN 242 339
FT	SITE 430 435
FT	SITE 453 458
FT	DISULFID 40 176
FT	DISULFID 45 108
FT	DISULFID 170 219
FT	DISULFID 278 323
FT	MOD_RES 432 432
FT	MOD_RES 455 455
FT	CARBOHYD 46 46
FT	CARBOHYD 167 167
FT	CARBOHYD 197 197
FT	CARBOHYD 216 216
FT	CARBOHYD 227 227
FT	CARBOHYD 237 237
FT	CARBOHYD 244 244
FT	CARBOHYD 262 262
FT	CARBOHYD 287 287
FT	CARBOHYD 294 294
FT	MUTAGEN 432 432
FT	MUTAGEN 455 455
FT	CONFLICT 2 2
FT	CONFLICT 24 24
FT	CONFLICT 102 102
FT	CONFLICT 234 234
FT	SEQUENCE 467 AA; 51889 MW; 54FD3E98950D2DEE CRC64;
CC	-----
CC	to as the immunoreceptor tyrosine-based inhibitor motif (ITIM).
CC	This motif is involved in downmodulation of cellular responses.
CC	The phosphorylated ITIM motif binds to the SH2 domain of
CC	PTPN6/SHP-1 and PTPN11/SHP-2.
CC	-!- PTM: Phosphorylated on tyrosine residues.
CC	-!- SIMILARITY: Belongs to the immunoglobulin superfamily, SIGLEC
CC	(sialic acid binding Ig-like lectin) family.
CC	-!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC	-!- SIMILARITY: Contains 2 immunoglobulin-like C2-type domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; AF317298; AAG38598.1; -.
DR	EMBL; AF329269; AAK49917.1; -.
DR	MED; MGI:1932475; Siglec5.
DR	GO; GO:0004872; F:receptor activity; IDA.
DR	GO; GO:0005057; F:receptor signaling protein activity; IPI.
DR	GO; GO:0007155; P:cell adhesion; IDA.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003598; Ig_c2.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00047; Ig; 2.
DR	SMART; SM00408; IGC2; 1.
DR	PROSITE; PS00835; IG_LIKE; 3.
DR	PROSITE; PS00290; IG_MHC; 1.
KW	Cell adhesion; Lectin; Transmembrane; Signal; Glycoprotein;
KW	Immunoglobulin domain; Repeat; Phosphorylation.
FT	SIGNAL 1 18
FT	CHAIN 19 467
FT	DOMAIN 19 353
FT	TRANSMEM 354 374
FT	DOMAIN 375 467
FT	DOMAIN 21 141
FT	DOMAIN 152 239
FT	DOMAIN 242 339
FT	DOMAIN 242 339
FT	SITE 430 435
FT	SITE 453 458
FT	DISULFID 40 176
FT	DISULFID 45 108
FT	DISULFID 170 219
FT	DISULFID 278 323
FT	MOD_RES 432 432
FT	MOD_RES 455 455
FT	CARBOHYD 46 46
FT	CARBOHYD 167 167
FT	CARBOHYD 197 197
FT	CARBOHYD 216 216
FT	CARBOHYD 227 227
FT	CARBOHYD 237 237
FT	CARBOHYD 244 244
FT	CARBOHYD 262 262
FT	CARBOHYD 287 287
FT	CARBOHYD 294 294
FT	MUTAGEN 432 432
FT	MUTAGEN 455 455
FT	CONFLICT 2 2
FT	CONFLICT 24 24
FT	CONFLICT 102 102
FT	CONFLICT 234 234
FT	SEQUENCE 467 AA; 51889 MW; 54FD3E98950D2DEE CRC64;
CC	-----
CC	to as the immunoreceptor tyrosine-based inhibitor motif (ITIM).
CC	This motif is involved in downmodulation of cellular responses.
CC	The phosphorylated ITIM motif binds to the SH2 domain of
CC	PTPN6/SHP-1 and PTPN11/SHP-2.
CC	-!- PTM:

```
Query Match          9.7%; Score 99; DB 1; Length 467;
Best Local Similarity 26.0%; Pred. No. 0.12;
Matches 34; Conservative 21; Mismatches 50; Indels 26; Gaps 4;
```

Cy	4	M L L I L I M V H P G S C A L W V O P P E I R T L - - - - - E G S A F L P C S F N A S Q G R L A I G S V T	54
Dd	1	M L L I L I L L W I K I G V E G Q N F V T I N V E R K V V V Q E G L C V L V P C N F S Y L K K L L T D W T D S	60
Cy	55	- - - - - W F R D E - - - - - V P Q K E V R N G T P P E R G L A P L A S S R F L H D H Q A E L H I R D V R G	100
Dd	61	D P V H G F W Y R E G T D R R K D S I V A T N N P - R K A V K E T R N P F F L G D P - - W R N D C S L N I R E I R K	117
Cy	101	H D A S I Y V C R V E	111
Dd	118	K D A G L Y F F R L E	128

RESULT 9  
BUTY\_HUMAN STANDARD; PRT; 526 AA.

ID	BUTY_HUMAN	STANDARD;	PRT;	526 AA.
AC	Q13410;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Butyrophilin precursor (BT) (Butyrophilin subfamily 1 member A1).			
GN	BTN1A1 OR BTN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
NCBI_TaxID=9606;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Breast;			
RX	MEDLINE=96201696; PubMed=8611614;			
RA	Taylor M.R., Peterson J.A., Ceriani R.L., Couto J.R.;			
RT	"Cloning and sequence analysis of human butyrophilin reveals a			
RT	potential receptor function."			
RL	Biochim. Biophys. Acta 1306:1-4(1996).			
CC	-I- FUNCTION: May function in the secretion of milk-fat droplets. It			
CC	may act as a specific membrane-associated receptor for the			
CC	association of cytoplasmic droplets with the apical plasma			
CC	membrane (By similarity).			
CC	-I- SUBUNIT: Seems to associate with xanthine dehydrogenase/oxidase			
CC	(By similarity).			
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-I- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG			
CC	family.			
CC	-I- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; U39576; AAC50489.1; -.			
DR	PIR; S70587; S70587.			
DR	GeneW; HGNC:1135; BTN1A1.			
DR	MIT; 601610; -.			
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.			
DR	GO; GO:0004872; F:receptor activity; TAS.			
DR	InterPro; IPRO01870; B302.			
DR	InterPro; IPRO07110; Ig-like.			
DR	InterPro; IPRO03596; Ig.			
DR	InterPro; IPRO06574; PRY.			
DR	InterPro; IPRO03877; SPRY_receptor.			
DR	Pfam; PF00047; Ig; 1.			
DR	Pfam; PF00622; SPRY; 1.			
DR	SMART; SM00406; IGV; 1.			
DR	SMART; SM00589; PRY; 1.			
DR	SMART; SM00449; SPRY; 1.			
DR	PROSITE; PS50835; IG_LIKE; 2.			
KW	Transmembrane; Glycoprotein; IMMUNOGLOBULIN domain; Signal; Repeat.			
FT	SIGNAL 1 26			
	BY SIMILARITY.			

DR SNART; SN00449; SPRY; 1.  
DR PROSITE; PS50835; IG\_LIKE; 2.  
KW Transmembrane; Glycoprotein; Immunoglobulin domain; Signal; Repeat.  
FT SIGNAL 1 26  
BY SIMILARITY.

**SIGNAL** 1



```

FT CHAIN 27 526 BUTYROPHILIN.
FT DOMAIN 27 242 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 243 269 POTENTIAL.
FT DOMAIN 270 526 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 138 IG-LIKE V-TYPE 1.
FT DOMAIN 148 234 IG-LIKE V-TYPE 2.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 526 AA; 59004 MW; E9EAC06F8DAF94D5 CRC64;

Query Match
Best Local Similarity 9.7%; Score 98.5; DB 1; Length 526;
Matches 44; Conservative 8; Mismatches 48; Indels 17; Gaps 8;

QY 4 MLLLLILMHP-GSCALWTSQPPE-IRTEGSAFLPC--SFNASQGRLAIGSVTWFRDE 59
DB 13 LTLILLQLPKLSDAPDFVIGPEPILAVVGDAELPCRLSPNASEHL---ELWFRKK 69
QY 60 VVPGKEV-RNG-----TFEFGRLAPLASSRFLHDHQAELHIRDVRGHDAIYVC 108
DB 70 VSPAVLHRDGRQEAQEPYRGR-ATLVQDGIAGRVA-LRIGRVVSDDGEYTC 124

RESULT 10
MOG RAT
ID MOG RAT STANDARD; PRT; 245 AA.
AC Q63345;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93085763; PubMed=1453482;
RA Gardiner M.V., Amiguet P., Linington C., Matthieu J.-M.;
RT "Myelin/oligodendrocyte glycoprotein is a unique member of the
RT immunoglobulin superfamily.";
RL J. Neurosci. Res. 33:177-187(1992).
RN [2]
RP SEQUENCE OF 28-245 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93376728; PubMed=8367453;
RA Pham-Dinh D., Mattei M.-G., Nusbaum J.-L., Roussel G.,
RA Pontarotti P., Roedel N., Mather I.H., Artzt K., Lindahl K.P.,
RA Dautigny A.;
RT "Myelin/oligodendrocyte glycoprotein is a member of a subset of the
RT immunoglobulin superfamily encoded within the major
RT histocompatibility complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7990-7994(1993).
RN [3]
RP STRUCTURE BY NMR OF 62-82.
RX MEDLINE=97354172; PubMed=9210466;
RA Albouze Abo S., Wilson J.C., Bernard C.C.A., von Itzstein M.;
RT "A conformational study of the human and rat encephalitogenic myelin
RT oligodendrocyte glycoprotein peptides 35-55.";
RL Eur. J. Biochem. 246:59-70(1997).
CC -1- FUNCTION: Minor component of the myelin sheath. May be involved in
CC completion and/or maintenance of the myelin sheath and in cell-
CC cell communication.
CC -1- SUBUNIT: May form homodimers.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Found exclusively in the CNS, where it is
CC localized on the surface of myelin and oligodendrocyte cytoplasmic
CC membranes.
CC -1- DEVELOPMENTAL STAGE: A PEAK OF EXPRESSION HAS BEEN OBSERVED
CC BETWEEN POSTNATAL DAYS 15 AND 25, COINCIDENT WITH THE PERIOD OF
CC ACTIVE MYELINATION.
CC -1- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG

```

```

CC family.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -1- CAUTION: Do not confuse myelin-oligodendrocyte glycoprotein (MOG)
CC with oligodendrocyte-myelin glycoprotein (OMG).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M99485; AAA41628.1; -
CC EMBL; L21995; AAF74786.1; -
CC PIR; B47712; B47712.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00406; IGv; 1.
CC PROSITE; PS00835; IG_LIKE; 1.
KW Immunoglobulin domain; Glycoprotein; Myelin; Transmembrane; Signal.
FT SIGNAL 1 27
FT CHAIN 28 245 MYELIN-OLIGODENDROCYTE GLYCOPROTEIN.
FT DOMAIN 28 155 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 156 176 POTENTIAL.
FT DOMAIN 177 208 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 209 229 POTENTIAL.
FT DOMAIN 230 245 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 30 139 IG-LIKE.
FT DISULFID 51 125 POTENTIAL.
FT CARBOHYD 58 58 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 245 AA; 27881 MW; C97F8AD60D6A32B4 CRC64;

Query Match
Best Local Similarity 9.8%; Score 98; DB 1; Length 245;
Matches 39; Conservative 14; Mismatches 47; Indels 20; Gaps 7;

QY 1 MAMWLLILIMVHPGSCALWTSQPPEIRTEGSAFLPCSFNASQGRLAIG-SVTWFRDE 59
DB 14 LSLULLLQLSRYAGQPRV-IGFCHPIRALVGDEALPC--RISPGKNATGMEVGHYRS- 69
QY 60 VVPGKEV-----RNG-----TFEFGRLAPLASSRFLHDHQAELHIRDVRGHDAIYVC 108
DB 70 --PFSRVVHLRYNGQDQAEAPYRGRTELLKES--IGEGKVALRIQNVRESDEGGYTC 125

RESULT 11
MOG_MOUSE
ID MOG_MOUSE STANDARD; PRT; 246 AA.
AC Q61885; P70364; Q62003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myelin-oligodendrocyte glycoprotein precursor.
GN MOG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129; PubMed=7829100;
RX MEDLINE=95130110;
RA Daubas P., Pham-Dinh D., Dautigny A.;
RT "Structure and polymorphism of the mouse myelin/oligodendrocyte
RT glycoprotein gene.";
RL Genomics 23:36-41(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Gardiner M.V., Matthieu J.M.;
RT "Murine and human MOG are highly conserved: cDNA analysis.";
RL Trans. Am. Soc. Neurochem. 24:234-234(1993).

```



QY 108 CR-VEVLGLGVGTGNGTRLVW 127  
 Db 110 CQSFNTNQG-VFGGTKLTV 129

RESULT 13  
 BUTY\_MOUSE  
 ID BUTY\_MOUSE STANDARD; PRT; 524 AA.  
 AC Q62556; P97392;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Butyrophilin precursor (BT) (Butyrophilin subfamily 1 member A1).  
 GN BTNL1 OR BTN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Mammary gland;  
 RX MEDLINE=97148936; PubMed=8995761;  
 RA Ogg S.L., Komaragiri M.V.S., Mather I.H.;  
 RT "Structural organization and mammary-specific expression of the  
 RT butyrophilin gene";  
 RL Mamm. Genome 7:900-905 (1996).  
 RN [2]  
 RP SEQUENCE OF 39-487 FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=96125722; PubMed=8541302;  
 RA Ishii T., Aoki N., Noda A., Adachi T., Nakamura R., Matsuda T.;  
 RT "Carboxy-terminal cytoplasmic domain of mouse butyrophilin  
 RT specifically associates with a 150-kDa protein of mammary epithelial  
 RT cells and milk fat globule membrane";  
 RL Biochim. Biophys. Acta 1245:285-292 (1995).  
 CC -!- FUNCTION: May function in the secretion of milk-fat droplets. It  
 CC may act as a specific membrane-associated receptor for the  
 CC association of cytoplasmic droplets with the apical plasma  
 CC membrane (By similarity).  
 CC -!- SUBUNIT: Seems to associate with xanthine dehydrogenase/oxidase.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed in mammary tissue and secreted in  
 CC association with the milk-fat-globule membrane during lactation.  
 CC -!- DEVELOPMENTAL STAGE: Expression increases during the last half of  
 CC pregnancy and is maximal during lactation.  
 CC -!- SIMILARITY: Belongs to the immunoglobulin superfamily. BTN/MOG  
 CC family.  
 CC -!- SIMILARITY: Contains 2 immunoglobulin-like V-type domains.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U67065; AAB51034.1; -  
 CC EMBL; S80642; AAB35893.1; -  
 CC MGI; MGI:103118; Btntl1.  
 CC InterPro; IPR001870; B302.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003596; IgV.  
 CC InterPro; IPR006574; PRY.  
 CC InterPro; IPR003877; SPRY\_receptor.  
 CC Pfam; PF00047; Ig; 1.  
 CC Pfam; PF00622; SPRY; 1.  
 CC SMART; SM00406; IGV; 1.  
 CC SMART; SM00589; PRY; 1.  
 CC SMART; SM00449; SPRY; 1.  
 CC PROSITE; PS50835; IGLIKE; 2. Immunoglobulin domain; Signal; Repeat.  
 CC Transmembrane; Glycoprotein; Immunoglobulin domain; Signal; Repeat.  
 KW

FT SIGNAL 1 26  
 FT CHAIN 27 524  
 FT DOMAIN 27 524  
 FT TRANSMEM 248 268  
 FT DOMAIN 269 524  
 FT DOMAIN 29 141  
 FT DOMAIN 149 235  
 FT CARBOHYD 56 56  
 FT CARBOHYD 216 216  
 FT CONFLICT 46 46  
 FT CONFLICT 117 117  
 FT CONFLICT 191 191  
 FT CONFLICT 210 210  
 FT CONFLICT 363 363  
 FT CONFLICT 408 408  
 FT CONFLICT 413 414  
 FT CONFLICT 420 423  
 FT CONFLICT 492 509  
 FT  
 SQ SEQUENCE 524 AA; 58406 MW; 333F4DEZC7704480 CRC64;

Query Match 9.3%; Score 95; DB 1; Length 524;  
 Best Local Similarity 29.1%; Pred. No. 0.33;  
 Matches 55; Conservative 17; Mismatches 51; Indels 66; Gaps 12;

QY 21 VSOPPE-IRTEGSSAFPCSF--NASQGLAIGSVTFWRD-----EYVPGKE 65  
 Db 32 VTAPQEPVLAIVGSDAELTCGSPNASSEYM---ELLNFRQTRSTAVLLYRDQGEQGG 88  
 QY 66 VRNGTPEFRGRLAPLASSRFLHDHQAELHIRVGRHDASIYVCRVEVLGLGVGTGNGTRL 125  
 Db 89 M-----TEYRGR-ATLTAGLL-DGRATLLIRDVRVSDQGEYRC----- 125  
 QY 126 VVEKEHPOLGAGVLLLRAGFYAVSFSLVAVGS-----TVYQKCHCHMGTHCHSSDG 179  
 Db 126 -LFRDNDDFEAAVYLKVA-----AVGSDPQISMTVQNGE---MELECTSS-- 168  
 QY 180 PRGVIEPR 188  
 Db 169 --GWYEPQ 175

RESULT 14  
 MOG\_HUMAN  
 ID -MOG\_HUMAN STANDARD; PRT; 247 AA.  
 AC Q16553; Q00713; Q00714; Q00715; Q13054; Q13055; Q14855; Q92891;  
 AC Q92892; Q92893; Q92894; Q92895; Q93053; Q96KU9; Q96KV0; Q99605;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Myelin-oligodendrocyte glycoprotein precursor.  
 GN MOG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=95310943; PubMed=7790876;  
 RA Hilton A.A., Slavik A.J., Hilton D.J., Bernard C.C.A.;  
 RT "Characterization of cDNA and genomic clones encoding human myelin  
 RT oligodendrocyte glycoprotein";  
 RL J. Neurochem. 65:309-318 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95054056; PubMed=7964757;  
 RA Pham-Dinh D., Allinquant B., Ruberg M., della Gaspera B.,  
 RA Nusebaum J.-L., Dautigny A.;  
 RT "Characterization and expression of the cDNA coding for the human  
 RT myelin/oligodendrocyte glycoprotein";  
 RL J. Neurochem. 63:2353-2356 (1994).  
 RN [3]



QY 72 --EFGELAPLASSRFLHDHQAELHIEDVGRHDASIVVCEVEVLGLGVGTGNGTRLVNVEK 129  
Db 90 WQFGRGRVWVGDD--ISKHDSGIVIRNLDYINDGTFTCDVKRPPDVVGTSVDHLTVYD 146  
QY 130 EHPQLGAGTV-----LLRAGFYAVSFL 152  
Db 147 KIPVVGAGVWGAIIGTFLGILLIIVGGLYLFYI 181

Search completed: February 26, 2004, 12:10:46  
Job time : 18.1273 secs

RESULT 15  
MYPO\_HETFR STANDARD; PRT; 246 AA.  
AC P20938;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Myelin P0 protein precursor (Myelin protein zero) (Myelin peripheral  
protein) (MPP).  
OS Heterodontus francisci (Horn shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Heterodontoides; Heterodontiformes;  
OC Heterodontidae; Heterodontus.  
OX NCBI\_TaxID=7792;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=90040744; PubMed=2478717;  
RA Saavedra R.A., Fors L., Aebersold R.H., Arden B., Horvath S.,  
RA Sanders J., Hood L.;  
RT "The myelin proteins of the shark brain are similar to the myelin  
proteins of the mammalian peripheral nervous system.";  
RL J. Mol. Evol. 29:149-156(1989).  
CC -!- FUNCTION: Creation of an extracellular membrane face which guides  
the wrapping process and ultimately compacts adjacent lamellae.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Found only in peripheral nervous system  
Schwann cells.  
CC -!- PTM: N-GLYCAN IS SULFATED (BY SIMILARITY).  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
CC -!- SIMILARITY: Belongs to the myelin P0 protein family.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X16714; CAB37865.1; -.  
CC PIR; A32999; A32999.  
CC HSSP; P06907; INEU.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003536; Ig\_V.  
DR InterPro; IPR000920; Myelin\_P0.  
DR PRINTS; PR00213; MYELINP0.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
DR PROSITE; PS00568; MYELIN\_P0; 1.  
KW Myelin; Structural protein; Glycoprotein; Transmembrane;  
KW Phosphorylation; Immunoglobulin domain; Signal.  
FT SIGNAL 1 27  
FT CHAIN 28 246 MYELIN P0 PROTEIN.  
FT DOMAIN 28 150 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 151 178 POTENTIAL.  
FT DOMAIN 179 246 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 28 143 IG-LIKE V-TYPE.  
FT DISULFID 48 125 POTENTIAL.  
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (COMPLEX) (BY  
SIMILARITY).  
FT SEQUENCE 246 AA; 27335 MW; A776A9ED9D430FA0 CRC64;

Query Match 9.1%; Score 93; DB 1; Length 246;  
Best Local Similarity 23.9%; Pred. No. 0.23;  
Matches 37; Conservative 26; Mismatches 66; Indels 26; Gaps 5;  
QY 21 VSOPPEIRTEGSAFLPCGSFNASQGRLAIGSVTW-FRD-----EVVPGKEVNGTFP---- 71  
Db 30 VSTHNLKHTKVSQDVTLYCGFWSNEYVSDLTLSMRFRPDNRSDIISIFHYGNVGYPIEK 89

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 28.7268 Seconds  
(without alignments)  
636.214 Million cell updates/sec

Title: US-10-036-444-2

Perfect score: 1020

Sequence: 1 MAWLLILIMVHPGSCALW.....GTHCHSSDGRGVIPRCP 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*  
1: Pirl.\*  
2: Pirl.\*  
3: Pirl.\*  
4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	11.6	136	2 B45893	T-cell receptor al
2	117.5	11.5	132	2 A24402	T-cell receptor al
3	116	11.4	146	2 S26408	T-cell receptor be
4	115.5	11.3	232	2 S25756	Ig lambda chain -
5	111.5	10.9	226	2 A46477	membrane-bound imm
6	111	10.9	233	2 S25747	Ig lambda chain -
7	110	10.8	223	2 T09536	cytotoxic T-lympho
8	108.5	10.6	223	2 A29063	cytotoxic T-lympho
9	108	10.5	235	2 S25749	Ig lambda chain -
10	107.5	10.5	138	2 C27577	T-cell receptor al
11	106.5	10.4	131	2 E24092	T-cell receptor al
12	106.5	10.4	131	2 D24092	T-cell receptor al
13	106	10.4	132	1 RWSAV	T-cell receptor al
14	105	10.3	131	2 F45893	T-cell receptor al
15	104.5	10.2	132	2 S09713	Ig lambda chain V
16	103.5	10.1	223	2 T46896	CTLA-4 precursor
17	103	10.1	129	2 A42692	T-cell receptor al
18	102.5	10.0	110	2 E24092	T-cell receptor al
19	101.5	10.0	139	2 S36325	T-cell receptor de
20	101	9.9	120	2 I54487	T-cell receptor al
21	101	9.9	235	2 S14675	Ig lambda chain -
22	100.5	9.9	1694	2 S50065	sialoadhesin - mou
23	100	9.8	110	2 A24092	T-cell receptor al
24	99.5	9.8	271	2 A53268	T-cell receptor al
25	99	9.7	130	2 A31211	T-cell receptor al
26	98.5	9.7	128	2 S24319	Ig lambda chain pr
27	98.5	9.7	152	2 S21826	T-cell receptor be
28	98.5	9.7	236	2 S25746	Ig lambda chain -
29	98.5	9.7	247	2 A55717	myelin/oligodendro

30 98.5 9.7 526 2 S70587 butyrophilin precu  
31 96 9.4 131 1 I6HUBB Ig lambda chain pr  
32 96 9.4 131 2 D24402 T-cell receptor al  
33 95.5 9.4 140 2 PH0134 Ig lambda chain pr  
34 95.5 9.4 218 2 B47712 myelin/oligodendro  
35 94.5 9.3 108 2 JH0342 T-cell receptor al  
36 94.5 9.3 186 2 S08614 cytotoxic T-lympho  
37 93.5 9.2 247 2 S58394 myelin/oligodendro  
38 93 9.1 117 2 I68824 T-cell receptor al  
39 93 9.1 132 2 D45893 T-cell receptor al  
40 93 9.1 246 1 A32999 myelin P0 protein  
41 93 9.1 267 1 RMSC8 T-cell receptor al  
42 92.5 9.1 136 2 S42610 ARMLambda protein  
43 92.5 9.1 139 2 S36302 T-cell receptor de  
44 92.5 9.1 403 2 I52590 m33-B isoform - mo  
45 92 9.0 129 2 A30554 Ig lambda chain pr

## ALIGNMENTS

### RESULT 1

B45893  
T-cell receptor alpha chain precursor V region (BTA29) - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000  
C:Accession: B45893  
R:Fishiguro, N.; Tanaka, A.; Shinagawa, M.  
Immunogenetics 31, 57-60, 1990  
A:Title: Sequence analysis of bovine T-cell receptor alpha chain.  
A:Reference number: A45893; MUID:90129157; PMID:2137108  
A:Accession: B45893  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-136 <ISR>  
A:Cross-references: GB:D90011; NID:g217610; PIDN:BAAL4061.1; PID:g217611  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: T-cell receptor

Query Match 11.6%; Score 118; DB 2; Length 136;  
Best Local Similarity 27.9%; Pred. No. 0.0016;  
Matches 41; Conservative 22; Mismatches 42; Indels 42; Gaps 9;

QY 4 MLLTLLIMV-----HPGSCALWVSPPEITRTLEGSAFLPCSFNAGSQRLAIGSVTFW 56  
DB 8 LLVILLASVDLEQPAEHS-----AEPASLPVEGAAASLGCTYSNSLIY-----FTWY 57

QY 57 RDEVVPGKEVRNGTPEF-----RGLAPLASSRFLHDHQAELHIRDVRGHDASI 105

DB 58 RQ--YPGK-----GPEFLQVYANNKEGKFT--AQSNTKNGH-VSLIRDSPEPSDSAT 107

QY 106 YVCRVSVLGLGVGT-----GNGTRLWVE 128

DB 108 YLCVADVTITTAGTKLTFEGTRLIVK 134

### RESULT 2

A24402  
T-cell receptor alpha chain precursor V region (C5) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 23-Aug-1987 #sequence\_revision 23-Aug-1987 #text\_change 30-May-1997

C:Accession: A24402

R:Becker, D.M.; Patten, P.; Chien, Y.; Yokota, T.; Eshhar, Z.; Giedlin, M.; Gascoigne,

Nature 317, 430-434, 1985

A:Title: Variability and repertoire size of T-cell receptor V-alpha gene segments.

A:Reference number: A33368; MUID:86014379; PMID:2995827

A:Accession: A24402

A:Molecule type: mRNA

A:Residues: 1-132 <BEC>

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: T-cell receptor

Query Match 11.5%; Score 117.5; DB 2; Length 132;

```

Best Local Similarity 30.8%; Pred. No. 0.0018;
Matches 44; Conservative 17; Mismatches 47; Indels 35; Gaps 9;

Qy 12 VHGSCALW-----VSOPPE-IRTEGSSAFPCSFNASGRLAIGSVTWFRD 59
Db 3 LHVSLVFLWTLQGGVSSQKQVQSPESLIYPEGAMVSLNCSFSDS-----ASQSIWYQQH 58
Qy 60 VVQKEVR-----NCTPFRGRGLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEV 112
Db 59 --FGKPKALLISFSGNKK-EGRLTYVYNRASH---VSLHKDQSPDSAVYLCVAVR 112
Qy 113 LGLGVGT---GNGRLVVEKEHP 132
Db 113 SGANTGKLTGHTILRV---HP 132

RESULT 3
S26408
T-cell receptor beta chain V region (5.5) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S26408
R: Bowman, S.J.; Larchbury, J.S.
submitted to the EMBL Data Library, September 1992
A:Description: T cell receptor beta chain sequences from patients with rheumatoid arthritis
A:Reference number: S26408
A:Accession: S26408
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-146 <BOW>
A:Cross-references: EMBL:X68527; NID:g36172; PIDN:CAA48540.1; PID:g36173
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:35-112/Domain: immunoglobulin homology <IMM>

Query Match 11.4%; Score 116; DB 2; Length 146;
Best Local Similarity 26.5%; Pred. No. 0.0027;
Matches 39; Conservative 26; Mismatches 52; Indels 30; Gaps 7;

Qy 1 MAWMLLLILIMVHGSCALWVSOPPE-IRTEGSSAFPCSFNASGRLAIGSVTWFRD 59
Db 6 LCWLLCLL---GAGSVETGVTQSPFTLIKT-RGQVTLRCSSQSQGH-----NTVSWYQQ 56
Qy 59 EVWPG-----KEVRNGTPEFRGRGLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRV 110
Db 57 ALGQGPQFIPOYYREENG-----RGNFPFSGLOPNYSSELNVNALELDDSAVLICAS 112
Qy 111 EVLGLGVGT-----GNGRLVVEKE 130
Db 113 SFKGLGLPSRGYEQYFGPGLTVTVD 139

RESULT 4
S25756
Ig lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S25756
R:Combrato, G.; Klobbeck, H.G.
Eur. J. Immunol. 21, 1513-1522, 1991
A:Title: (Vlambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lambda
A:Reference number: S16439; MUID:91257162; PMID:1904362
A:Accession: S25756
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-232 <COM>
A:Cross-references: EMBL:X57821; NID:g33741; PIDN:CAA40958.1; PID:g33742
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:147-215/Domain: immunoglobulin homology <IMM>

Query Match 11.3%; Score 115.5; DB 2; Length 232;
Best Local Similarity 26.3%; Pred. No. 0.0049;

```

```

Matches 46; Conservative 24; Mismatches 62; Indels 43; Gaps 8;

Qy 1 MAWMLLLILIMVHGSCALWVSOPPEIRTEGSSAFPCSFNASGRLAIGSVTWFRD 58
Db 1 MAWTVLLGLLSHCTGVTSVLTQPPSVVAPKRTASITCGGN-----NIGSKSVHWYQQ 56
Qy 59 EVWPGK-----EVRNGTPE-FRGRGLAPLASSRFLHDHQAELHIRDVRGHDASIYV 107
Db 57 K--FQAPLVVYVDDSDRPSGIPERFSGSNS-----GNTATLTISRVEAGDEADYY 105
Qy 108 CRVEVLGCVGTGNGTRLV-----EKEHFGLGAGTIVLLRAGFY 147
Db 106 CQVMDSSDVVFGGTKTLTVLGQPKAAPSVTLFPPSSEELQANKATLVCLISDFY 160

RESULT 5
A46477
membrane-bound immunoglobulin Ig-alpha chain precursor - human
N:Alternate names: B-cell antigen receptor complex alpha chain CD79a; IgM-alpha; immunoglobulin heavy chain 1
C:Species: Homo sapiens (man)
C>Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: I54539; A46477; A49135; A46479; I54496; I57851; S51113
R:Hashimoto, S.; Mohrenweiser, H.W.; Gregersen, P.K.; Chiorazzi, N.
Immunogenetics 40, 287-295, 1994
A:Title: Chromosomal localization, genomic structure, and allelic polymorphism of the human B-cell antigen receptor complex alpha chain CD79a
A:Reference number: I54539; MUID:94364636; PMID:7916003
A:Accession: I54539
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-226 <HAI>
A:Cross-references: GB:L32754; NID:g600875; PIDN:AAA57274.1; PID:g602606
R:Yu, L.M.; Chang, T.W.
J. Immunol. 148, 633-637, 1992
A:Title: Human mb-1 gene: complete cDNA sequence and its expression in B cells bearing the B-cell antigen receptor complex
A:Reference number: A46477; MUID:92105765; PMID:1729378
A:Accession: A46477
A:Molecule type: mRNA
A:Residues: 1-226 <YU>
A:Cross-references: GB:S75217; NID:g241773; PIDN:AA20812.1; PID:g241774
A:Note: sequence extracted from NCBI backbone (NCBIN:75217, NCBIF:75218)
R:LeDuc, I.; Preud'homme, J.L.; Cogne, M.
Clin. Exp. Immunol. 90, 141-146, 1992
A:Title: Structure and expression of the mb-1 transcript in human lymphoid cells.
A:Reference number: A49135; MUID:93009083; PMID:1395095
A:Accession: A49135
A:Molecule type: mRNA
A:Residues: 1-226 <LE>
A:Cross-references: GB:S46706; NID:g257140; PIDN:AA23558.1; PID:g257141
A:Experimental source: Ly66 cell line
A:Note: sequence extracted from NCBI backbone (NCBIN:115213, NCBIF:115214)
R:Ha, H.J.; Kubagawa, H.; Burrows, P.D.
J. Immunol. 148, 1526-1531, 1992
A:Title: Molecular cloning and expression pattern of a human gene homologous to the murine B-cell antigen receptor complex alpha chain CD79a
A:Reference number: A46479; MUID:92168394; PMID:1538135
A:Accession: A46479
A:Molecule type: mRNA
A:Residues: 1-226 <HAI>
A:Cross-references: GB:U05259; NID:g452561; PIDN:AAA20495.1; PID:g521105; NCBIF:85181
R:Flaswinkel, H.; Reth, M.
Immunogenetics 36, 268-269, 1992
A:Title: Molecular cloning of the Ig-alpha subunit of the human B-cell antigen receptor complex
A:Reference number: I54496; MUID:92347937; PMID:1639443
A:Accession: I54496
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-68, 1' 70-226 <FLA>
A:Cross-references: GB:M74721; NID:g337419; PIDN:AAA60270.1; PID:g337420
R:Hashimoto, S.; Chiorazzi, N.; Gregersen, P.K.
Mol. Immunol. 32, 651-659, 1995
A:Title: Alternative splicing of CD79a (Ig-alpha/mb-1) and CD79b (Ig-beta/B29) RNA transcripts
A:Reference number: I57851; MUID:95371688; PMID:7643857
A:Accession: I57851

```



A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-88, 'E', 128-226 <HAS3>  
A:Cross-references: GB:S79248; NID:g1087006; PIDN:AA60653.1; PID:g1087007  
R:Koyama M., Nakamura, T.  
Submitted to the EMBL data Library, December 1994  
A:Reference number: S5112  
A:Accession: S5113  
A:Molecule type: mRNA  
A:Residues: 1-88, 'E', 128-226 <KOV>  
A:Cross-references: EMBL:X83540; NID:g620080; PIDN:CAA59523.1; PID:g620081  
C:Genetics:  
A:Gene: GDB:TGA; mb-1; CD79a  
A:Cross-references: GDB:133778; OMIM:600352  
A:Map position: 19q13.2-19q13.2  
A:Introns: 27/1; 127/1; 166/3; 190/1  
C:Complex: disulfide-linked heterodimer with Ig-beta; heterodimer associates non-covalently  
C:Keywords: alternative splicing; disulfide bond; glycoprotein; immunoglobulin; surface  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-226/Product: membrane-bound immunoglobulin Ig-alpha chain #status predicted <MAT>  
F:33-143/Domain: extracellular #status predicted <EXT>  
F:144-165/Domain: transmembrane #status predicted <TM>  
F:166-226/Domain: intracellular #status predicted <CVT>  
F:57,63,73,89,97,112/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.9%; Score 111.5; DB 2; Length 226;  
Best Local Similarity 28.0%; Pred. No. 0.011; Mismatches 24; Conservative 47; Indels 70; Gaps 8;  
Matches 47; Conservative 24; Mismatches 24; Indels 70; Gaps 8;

QY 4 MLLLLIMVHPCALWVSQ-PPEIRTLGGSAFLPCSFNASQGLAIGSVTWRDEVVP 62  
DB 18 LFLLSAVLPGCCQALWVHKVPSALVSLGDAHPQCPNSNN-----ANVTWR--VJH 71  
QY 63 GKEVRNGT--PFRGLAPLASSRFLHDHQLHVRDVRGHDSIYVCRVEVLGVGTG 120  
DB 72 G-----NYTPPEFLGP-----GDPNGLLIQNVKSHGGIYVCRVQEGNESYQOS 118  
QY 121 NGTRLVVKEHFQ-----LGAGTV-LLLRAGFVAVSFLSVAGSTVYVQ 163  
DB 119 CGTYLVRGPPRPFLDMGEGTNRITIAEGILLFCAVPGTLLFR 166

RESULT 6  
S25747  
Ig lambda chain - human  
C:Species: Homo sapiens (man)  
C>Date: 22-Nov-1993 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
C:Accession: S25747  
R:Combarato, G.; Klobbeck, H.G.  
Eur. J. Immunol. 21, 1513-1522, 1991  
A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam  
A:Reference number: S16439; MUID:91257162; PMID:1904362  
A:Accession: S25747  
A>Status: preliminary; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-233 <COM>  
A:Cross-references: EMBL:X57812; NID:g33723; PIDN:CAA40949.1; PID:g33724  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
F:148-216/Domain: immunoglobulin homology <IMV>

Query Match 10.9%; Score 111; DB 2; Length 233;  
Best Local Similarity 28.0%; Pred. No. 0.012; Mismatches 24; Conservative 49; Indels 60; Gaps 10;  
Matches 49; Conservative 24; Mismatches 60; Indels 60; Gaps 10;

QY 1 MAWMLLLIMVH-PGSCALWV-SOPPEIRTLGGSAFLPCSFNASQGLAIGSVTWR 57  
DB 1 MANTVLLGLLHSGTGSVTSYVLTQPPSVSVAPGVAARITCG-----GINTASVHHYQ 55  
QY 58 DE-----VPGKEVR-NGTPE-FRGLAPLASSRFLHDHQLHVRDVRGHDSIYV 108  
DB 56 QKPGQAPVLVWYGDSDRPSGIPERFSGNS-----GNTATLNISRVEAGDEAAAYC 106

QY 109 RV-EVLGLGVGTGNGTRLVV-----EKEHPQLGAGTVLLLRAGFY 147  
DB 107 QVWDSSDHVVFGGGKLTVLGQPKAPSVTLFPPPSSEELQANKATIVCLISDFY 161

RESULT 7  
T09536  
cytotoxic T-lymphocyte protein 4 - human  
C:Species: Homo sapiens (man)  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T09536  
R:Harper, K.; Balzano, C.; Rouvier, E.; Mattei, M.G.; Luciani, M.F.; Golstein, P.  
J. Immunol. 147, 1037-1044, 1991  
A:Title: CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mou  
A:Reference number: I49584; MUID:91318145; PMID:1713603  
A:Accession: T09536  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-223 <HAR>  
A:Cross-references: EMBL:LI5006; NID:g291928; PIDN:AA859385.1; PID:g291929  
C:Genetics:  
A:Gene: CTLA4  
A:Map position: 2q33  
A:Superfamily: T-cell surface glycoprotein CD28; immunoglobulin homology  
C:Keywords: T-cell; transmembrane protein

Query Match 10.8%; Score 110; DB 2; Length 223;  
Best Local Similarity 28.0%; Pred. No. 0.015; Mismatches 22; Conservative 47; Indels 24; Gaps 6;  
Matches 47; Conservative 22; Mismatches 22; Indels 24; Gaps 6;

QY 6 LLLILVHVGSC-ALWVSOPPEIRTLGGSAFLPCSFNASQGLAIGSVTWRDEVVP 64  
DB 24 LLEFLFIPIVFCAMHVAQPAVVLASSRGIAISFVCEY-ASPGKATEVRVTILRQDSQVT 82  
QY 65 EVANGPEPEFRGLAPLASSRFLHD-----HQAELHVRDVRGHDSIYVCRVEVL-- 113  
DB 83 EVCAATVTGNETL-----FLDSDICTGTSSGNQVNLTIQGLRDMDTGLYICKVELMYP 136  
QY 114 -GLGVGTGNGTRLVVEKEHPQLGAGTVLLLR-----GFYAVSFLSVAV 156  
DB 137 PPTLIGNGTQIYVIDPEPCPDSDFLMLAAVSSGLFFYSFLTAV 184

RESULT 8  
A29063  
Cytotoxic T-lymphocyte protein 4 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 05-Nov-1999  
C:Accession: A29063; I49622  
R:Brust, J.F.; Denizot, F.; Luciani, M.F.; Roux-Dosseto, M.; Suzan, M.; Mattei, M.G.;  
Nature 328, 267-270, 1987  
A:Title: A new member of the immunoglobulin superfamily--CTLA-4.  
A:Reference number: A29063; MUID:8758259; PMID:3496540  
A:Accession: A29063  
A:Molecule type: mRNA  
A:Residues: 1-223 <BRU>  
A:Cross-references: GB:X05719; NID:g50592; PIDN:CAA2191.1; PID:g50593  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
R:Harper, K.; Balzano, C.; Rouvier, E.; Mattei, M.G.; Luciani, M.F.; Golstein, P.  
J. Immunol. 147, 1037-1044, 1991  
A:Title: CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mou  
A:Reference number: I49584; MUID:91318145; PMID:1713603  
A:Accession: I49622  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-36 <RES>  
A:Cross-references: GB:M74362; NID:g192833; PIDN:AAA37489.1; PID:g553903  
C:Genetics:  
A:Gene: Ctla-4  
A:Map position: 1, band C  
C:Superfamily: T-cell surface glycoprotein CD28; immunoglobulin homology  
C:Keywords: transmembrane protein

Query Match 10.6%; Score 108.5; DB 2; Length 223;  
Matches 49; Conservative 24; Mismatches 60; Indels 60; Gaps 10;



```
Best Local Similarity 26.9%; Pred. No. 0.02;
Matches 45; Conservative 26; Mismatches 79; Indels 17; Gaps 6;

QY 3 WMILLILWVPGSCALWVSOPPIRELEGSAPLPCSFNASQGRLAIGSVTWFRD-----58
DB 22 FVALLTLFIFVFSEAIQVTPQPSVVLASSHGVSFPCPSFHTDEV-RVTVLRQTNQD 80
QY 59 --EVVPGKEVRNGTPFERGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVL---113
DB 81 MTEVCATTFTEKNTVGFLD--YPCSGTFF-NESRVNLTIOGLRAVDGTGLYCKVELMYP 137
QY 114 GLGVGTGNGTFLVKEKHPQLGAGTVL-----LLRAGFYAVSFSLVAV 156
DB 138 PYFVGMGNGTQIYVIDPECPDSDFLWILWAVSLGLFFSYFLVSAV 184

RESULT 9
S25749
Ig lambda chain - human
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C:Accession: S25749
R:Combrinato, G.; Klobbeck, H.G.
A:Title: V(lambda) and J(lambda)-C(lambda) gene segments of the human immunoglobulin lam
A:Reference number: S16439; MUID:91257162; PMID:1904382
A:Accession: S25749
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Cross-references: EMBL:X57814; NID:933727; PIDN:CAA40951.1; PID:933728
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:150-218/Domain: immunoglobulin homology <IMM>

Query Match 10.6%; Score 108; DB 2; Length 235;
Best Local Similarity 25.1%; Pred. No. 0.023;
Matches 45; Conservative 26; Mismatches 60; Indels 48; Gaps 10;

QY 1 MAMWMLLILIMVHFGS---CALWVSOPPEIRTELEGSAPLPCSFNASQGRLAIGSVTWFR 57
DB 1 MAMWMLLILGLLAY-GSGVDGQTVTVTPQPSFMSFGGTVTTCGLSSGSGSTY-SPSWYQ 58
QY 58 DEVVPGKEVR-----NGTPE-FRGLAPLASSRFLHDHQAELHIRDVRGHDAIY 106
DB 59 Q--TFGQAPRTLMYNTTRSSGVDFRSGSIL-----GNKSALTITGAQEDSDY 107
QY 107 VCRVEVLGLVGT---GNGTFLV-----EKEHPQLGAGTVLLRAGFY 147
DB 108 YC-----VLYMGRGIVFGGKTLTVLGQPKAAPSVTLFPPPSSEELQANKATLVCLISDFY 163

RESULT 10
C27577
T-cell receptor alpha chain V region (5/10-20D) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-May-1997
C:Accession: C27577; C27557
R:Iwamoto, A.; Ohashi, P.S.; Picher, H.; Walker, C.L.; Michalopoulos, E.E.; Rupp, F.; H
J. Exp. Med. 165, 591-600, 1987
A:Title: T cell receptor variable gene usage in a specific cytotoxic T cell response. P
A:Reference number: A27557; MUID:87139812; PMID:3493320
A:Accession: C27577
A:Molecule type: mRNA
A:Residues: 1-138 <1WA>
A:Accession: C27557
A:Molecule type: mRNA
A:Residues: 22-138 <1W2>
A:Cross-references: EMBL:X05733
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match 10.5%; Score 107.5; DB 2; Length 138;
```

```
Best Local Similarity 31.9%; Pred. No. 0.015;
Matches 38; Conservative 17; Mismatches 45; Indels 19; Gaps 7;

QY 21 VSQPE-IRTELEGSAPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVR-----NGTPE 72
DB 24 VQSPESLIVPEGATMCTSLNCTFSDASQY-----FATYRQH--SGKAPKALMSIFSNBKE 77
QY 73 FRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVL-GLVGTGNGTFLVVEKE 130
DB 78 -EGRTIHLNKSLH---FSLHIRDSPSDSALYCAVMTGGYKVVFGSGTRLVSPD 132

RESULT 11
E24092
T-cell receptor alpha chain precursor V region (4.C3) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 21-Jan-2000
C:Accession: E24092
R:Finck, P.J.; Mattis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A:Title: Correlations between T-cell specificity and the structure of the antigen recep
A:Reference number: A93380; MUID:86230843; PMID:3012351
A:Accession: E24092
A:Molecule type: mRNA
A:Residues: 1-131 <FIN>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:21-131/Product: T-cell receptor alpha chain V region 4.C3 #status predicted <MAT>
F:36-109/Domain: immunoglobulin homology <IMM>

Query Match 10.4%; Score 106.5; DB 2; Length 131;
Best Local Similarity 29.6%; Pred. No. 0.017;
Matches 37; Conservative 19; Mismatches 42; Indels 27; Gaps 7;

QY 21 VSQPEIRTL-EGSSAPLPCSFNASQGRLAIGSVTWFR-----RDSVWPGKEVRNGTPEFRG 75
DB 24 VEQSPSALSLEHTGSALRCNFTTT-----MRAVQWFRKNSRGLINFLVLAQTKE-NG 77
QY 76 RLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLVGTGNGTFLVVEKEHPQLG 135
DB 78 RLKSAFDSK---BRYSTLHIRDALQEDSGTYFCAE-----STSSQGLV-----FG 121
QY 136 AGTVL 140
DB 122 QGTVL 126

RESULT 12
D24092
T-cell receptor alpha chain precursor V region (B10) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Nov-1988 #sequence_revision 01-Dec-2000 #text_change 11-Jan-2002
C:Accession: D24092; S03507
R:Finck, P.J.; Mattis, L.A.; McElligott, D.L.; Bookman, M.; Hedrick, S.M.
Nature 321, 219-226, 1986
A:Title: Correlations between T-cell specificity and the structure of the antigen recep
A:Reference number: A93380; MUID:86230843; PMID:3012351
A:Accession: D24092
A:Molecule type: mRNA
A:Residues: 1-131 <FIN>
R:Winkto, A.; Mjolsness, S.; Hood, L.
Nature 316, 832-836, 1985
A:Title: Genomic organization of the genes encoding mouse T-cell receptor alpha-chain.
A:Reference number: S03503; MUID:85296332; PMID:2993908
A:Accession: S03507
A:Molecule type: DNA
A:Residues: 111-130 <WIN>
A:Cross-references: EMBL:X03057; NID:954519; PIDN:CAA6864.1; PID:gl334131
A>Note: this sequence was determined from the germline gene
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor
F:1-20/Domain: signal sequence #status predicted <SIG>
```

F/21-131/Product: T-cell receptor alpha chain V region B10 #status predicted <MAT>  
F/36-109/Domain: immunoglobulin homology <IMM>

Query Match 10.4%; Score 106.5; DB 2; Length 131;  
Best Local Similarity 29.6%; Pred. No. 0.017;  
Matches 37; Conservative 19; Mismatches 42; Indels 27; Gaps 7;

QY 21 VSOPPEIRTL-EGSSAFPCSFNASQGRLAIGSVTWTF-----RDEVVPGKEVRNGTPEFRG 75  
Db 24 VEQSPSALSHEGTGSALRCNFTTT-----MRAVQFRKNRSGSLNLFVLASGTKE-NG 77

QY 76 RLAPLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGVGTGTGTLVVEKEHPQLG 135  
Db 78 RLKSAFDK---ERYSTLHIDAQLEDSGTIFYCAAE-----ATSSGQKLIV-----FG 121

QY 136 AGTVL 140

Db 122 QGTIL 126

#### RESULT 13

RWMSAV

T-cell receptor alpha chain precursor V region (2B4) - mouse

C/Species: Mus musculus (house mouse)

C/Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 30-May-1997

C/Accession: A02015

R/Chien, Y.; Becker, D.M.; Lindsten, T.; Okamura, M.; Cohen, D.I.; Davis, M.M.

Nature 312, 31-35, 1984

A/Title: A third type of murine T-cell receptor gene.

A/Reference number: A93344; MUID:85036634; PMID:6548551

A/Accession: A02015

A/Molecule type: mRNA

A/Residues: 1-132 <CHI>

A/Experimental source: hybridoma 2B4, clone T711

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: glycoprotein; heterotetramer; receptor; T-cell

F/1-20/Domain: signal sequence #status predicted <SIG>

F/21-132/Product: T-cell receptor alpha chain V region (2B4) #status predicted <MAT>

F/21-113/Region: V segment

F/114-117/Region: D segment

F/118-132/Region: J segment

F/42/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 106; DB 1; Length 132;  
Best Local Similarity 25.8%; Pred. No. 0.019;  
Matches 40; Conservative 22; Mismatches 45; Indels 48; Gaps 9;

QY 3 WMLLLILIMVHFGSCALWVS-----QPPEIRTEGSAFLPCSFNASQGRLAIGSVTW 55  
Db 12 WLLL-----NWNQONVOQSPESLIVPEGARTSLNCTFSDASQY----FW 55

QY 56 FRDEVVPGKEVR-----NGTPEFRGRPLAPLASSRFLHDHQAELHIRDVRGHDASIVY 108  
Db 56 YRQH--SGRKAPKALMSIFSGEKE--SGRTIHLNKASLH---FSLHIRDSQPSDSALYIC 109

QY 109 RVEVLGVGTGTGTLVVEKEHPQLGAGTVLLLR 143

Db 110 AVTLYG-----GSGNKL-----FGTGTLLSVK 132

#### RESULT 14

F45893

T-cell receptor alpha chain precursor V region (BTA25) - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000

C/Accession: F45893

R/Ishiguro, N.; Tanaka, A.; Shinagawa, M.

Immunogenetics 31, 57-60, 1990

A/Title: Sequence analysis of bovine T-cell receptor alpha chain.

A/Reference number: A45893; MUID:90129157; PMID:2137108

A/Accession: F45893

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-131 <ISH>

A/Cross-references: GB:D50015; NID:G217618; PIDN:BAA14065.1; PID:G217619

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: T-cell receptor

F/36-109/Domain: immunoglobulin homology <IMM>

Query Match 10.3%; Score 105; DB 2; Length 131;  
Best Local Similarity 30.3%; Pred. No. 0.023;  
Matches 36; Conservative 23; Mismatches 34; Indels 26; Gaps 8;

QY 21 VSOPPEIRTL-EGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKE-----VNGTPEF 73  
Db 24 VEQSPSVLSLQEGANSTLRNFSDT-----VDSYQVFQON--PGCALTTLFFIASGTKK- 75

QY 74 RGRPLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGVGT-----GNGTRLV 127  
Db 76 NERMSSTVNSK---ERYSTLHTASQLEDAATLYCAVD---LQSGNQFLVFGKGRRLAV 128

#### RESULT 15

S09713

IG lambda chain V region - human

C/Species: Homo sapiens (man)

C/Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000

C/Accession: S09713

R/Rhodes-Jones, N.C.; Bye, J.M.; Beale, D.; Coadwell, J.

Biochem. J. 268, 135-140, 1990

A/Title: Nucleotide sequences and three-dimensional modelling of the VH and VL domains

A/Reference number: S09710; MUID:90262535; PMID:2111699

A/Accession: S09713

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-132 <HUG>

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: heterotetramer; immunoglobulin

F/34-110/Domain: immunoglobulin homology <IMM>

Query Match 10.2%; Score 104.5; DB 2; Length 132;  
Best Local Similarity 29.8%; Pred. No. 0.026;  
Matches 42; Conservative 20; Mismatches 54; Indels 25; Gaps 8;

QY 1 MAMVLLILIMVH-PGSCALWV-SQPPEIRTEGSAFLPCSFNASQGRLAIGSVTWFRD 58  
Db 1 MAMTVLLGLLSHCTGTSVTLTPPSVSAPQGTARITC--GGTSNNIGRKSVHWYQQ 58

QY 59 E-----VVPGEYR-NGTPE-FRGLAPLASSRFLHDHQAELHIRDVRGHDASIVY 109  
Db 59 KPGQAPVLVYVGASDRPSGIPERFSGSNS-----GNTATLTISRVAAGDEADYYCQ 109

QY 110 V---EVILGLGVGTGNGTRLVV 127

Db 110 WVDSSSAHPGVFGGGLKLV 130

Search completed: February 26, 2004, 12:14:49  
Job time : 30.7268 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 117.931 Seconds  
(without alignments)  
455.215 Million cell updates/sec

Title: US-10-036-444-2

Perfect score: 1020

Sequence: 1 NAWMLLLILIMVHPSGALW.....GTHCHSSDGRGVIPEPRCP 190

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	190	AAE02769	Aae02769 Human NKp
2	1018	99.8	190	AAY06401	Aay06401 Human B-c
3	876	85.9	201	AAY06403	Aay06403 Human B-c
4	859	84.2	177	AAY06402	Aay06402 Human B-c
5	713	69.9	135	AAE19109	Aae19109 Human NKp
6	713	69.9	369	AAE19110	Aae19110 Human NKp
7	632	62.0	120	AAE02771	Aae02771 Human NKp
8	204	20.0	33	AAE02773	Aae02773 Human NKp
9	123.5	12.1	246	AAE20273	Aae20273 Human lun
10	119.5	11.7	232	ABU08020	Abu08020 Monoclonal
11	118.5	11.6	139	AAU75565	Aau75565 Murine T
12	118.5	11.6	212	AAE60134	Aae60134 CTLA4 rec
13	118.5	11.6	212	AAU77642	Aar77642 Full leng
14	118.5	11.6	212	AAU43479	Aay43479 Amino aci
15	118.5	11.6	212	AAW81584	Aaw81584 Human CTL
16	118.5	11.6	212	ABU78106	Abu78106 Amino aci
17	118.5	11.6	212	AAU75124	Aau75124 Human CTL
18	118.5	11.6	212	AAU75132	Aau75132 Human CTL
19	118.5	11.6	212	ABU56715	Abu56715 Human CTL
20	118.5	11.6	212	ADU89013	Adu89013 CTLA4 rec
21	116.5	11.4	212	AAU75126	Aau75126 Human CTL
22	116	11.4	211	AAW87560	Aaw87560 Human CTL
23	115.5	11.3	212	AAU75127	Aau75127 Human CTL
24	115.5	11.3	212	AAU75131	Aau75131 Human CTL
25	115.5	11.3	212	AAU75130	Aau75130 Human CTL

26	114.5	11.2	212	5	AAU75133	Aau75133 Human CTL
27	112	11.0	222	6	ABP5948	Abp5948 Human CTL
28	112	11.0	223	3	AAU15129	Aau15129 Human CTL
29	112	11.0	223	5	AAU74509	Aau74509 Human cyt
30	112	11.0	223	5	AAU74508	Aau74508 Human cyt
31	112	11.0	235	4	AAU64474	Aau64474 Human typ
32	111.5	10.9	212	2	AAU31040	Aau31040 Human CTL
33	111.5	10.9	212	5	AAU75129	Aau75129 Human CTL
34	111.5	10.9	226	2	AAU24112	Aau24112 Human mb-
35	111.5	10.9	226	3	AAU52553	Aau52553 Human MB-
36	111.5	10.9	226	6	ABG74347	Abg74347 Human Iga
37	111.5	10.9	226	7	ADD67577	Add67577 Mouse CTL
38	110.5	10.8	223	5	ABB79935	Abb79935 Mouse CTL
39	110	10.8	223	4	AAU25111	Aau25111 Soluble h
40	110	10.8	223	4	AAU00687	Aau00687 Human CTL
41	110	10.8	223	5	ABB79934	Abb79934 Human CTL
42	110	10.8	223	7	ADA50039	Ada50039 Human cyt
43	110	10.8	223	7	ADC78847	Adc78847 Human PRO
44	110	10.8	223	7	ADD25540	Add25540 Binding d
45	110	10.8	235	4	AAG64476	Aag64476 Human typ

ALIGNMENTS

RESULT 1

AAE02769  
ID AAE02769 standard; protein; 190 AA.

XX AC AAE02769;

XX DT 06-AUG-2001 (first entry)

XX DE Human NKp30 receptor.

XX KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;

KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
KW therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 1..18

XX FT Protein /label= Signal\_peptide

XX FT Region /label= Mature\_Nkp30\_receptor\_protein

XX FT Region /label= Extracellular\_region

XX FT Modified-site 42 /note= "Forms an immunoglobulin (Ig) V-like domain"

XX FT Modified-site 121 /note= "N-glycosylation site"

XX FT Region 139..157 /note= "N-glycosylation site"

XX FT Region 158..190 /label= Transmembrane\_region

XX FT Region /label= Intracellular\_region

XX FN WO200136630-A2.

XX PD 25-MAY-2001.

XX PF 15-NOV-2000; 2000MO-EP011697.

XX PR 15-NOV-1999; 99CA-02288307.

XX PR 15-NOV-1999; 99US-00440514.

XX PA (INNA-) INNATE PHARMA SAS.

XX PA (UYGE-) UNIV GENOVA.

XX PI Moretta A, Bottino C, Biassoni R;

```

DR WPI; 2001-329221/34.
DR N-PSDB; AAD06564.
XX
XX Novel compound, useful for detection and/or quantifying the presence of
XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.
XX
XX Claim 1; Fig 7B; 83pp; English.
XX
XX The invention relates to human Nkp30 receptor and its corresponding cDNA
XX molecule which is involved in natural cytotoxicity mediated by natural
XX killer (NK) cells and antibodies that identify the same. Nkp30 receptor
XX is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively
XX expressed on the surface of human mature NK cells. Nkp30 and its cDNA are
XX useful for detecting and/or quantifying the presence of NK cells in a
XX biological sample. The invention also provides kits for detecting and/or
XX quantifying the presence of NK cells, for the selective removal of NK
XX cells from a biological sample, for the positive and selective
XX purification of NK cells from a biological sample and for the in vitro
XX stimulation of NK cell cytotoxicity. The invention further provides a
XX pharmaceutical composition which is used as a drug for grafting
XX enhancement, graft versus host (GvH) inhibition, stimulation of graft
XX versus tumour (GvT) and especially graft versus leukaemia (GvL), and for
XX the prevention, palliation and/or therapy of solid or liquid tumours,
XX such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or
XX microorganism, notably viral infection. Nkp30 antibodies are useful for
XX identifying Nkp30 natural ligands and allow assessment of the level of
XX surface Nkp30 ligand expressed on an NK-susceptible target cell and the
XX comparison of this level to the standard physiological one. Hence Nkp30
XX antibodies are useful in the diagnosis of tumours or of infection. The
XX present sequence is human Nkp30 receptor
XX
XX Sequence 190 AA;
XX
XX Query Match 100.0%; Score 1020; DB 4; Length 190;
XX Best Local Similarity 100.0%; Pred. No. 6.6e-92;
XX Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAWLLILLIMVPGSCALWVSPPEIRTLGSSAFPLPCSFNASQGRLAIGSVTFRDEV 60
DB 1 MAWLLILLIMVPGSCALWVSPPEIRTLGSSAFPLPCSFNASQGRLAIGSVTFRDEV 60
QY 61 VPGKEVNGTPEFRGRPLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
DB 61 VPGKEVNGTPEFRGRPLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
QY 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYVYQKCHCHMGTHCHSSDGP 180
DB 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYVYQKCHCHMGTHCHSSDGP 180
QY 181 RGVIPEPRCP 190
DB 181 RGVIPEPRCP 190
XX
XX RESULT 2
XX ID AAY06401 standard; protein; 190 AA.
XX AC AAY06401;
XX XX
XX 20-SEP-1999 (first entry)
XX
XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.
XX
XX BMOG; B-cell myelin oligodendrocyte glycoprotein; human;
XX signal transduction; immunomodulator; antiinflammatory;
XX autoimmune disease; inflammation; gene therapy; diagnosis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..12
XX /note= "leader peptide"
XX

```

---

```

FT Protein 13. .190
FT Modified-site /note= "mature protein"
FT 42
FT Modified-site /note= "N-glycosylated"
FT 68
FT Modified-site /note= "N-glycosylated"
FT 121
FT Modified-site /note= "N-glycosylated"
FT 139..162
FT Domain /note= "transmembrane domain"
FT Peptide 166..190
FT /note= "alternatively spliced C-terminal end"
XX
XX WO9923867-A2.
XX 20-MAY-1999.
XX
XX 05-NOV-1998; 98WO-US023826.
XX
XX 07-NOV-1997; 97US-0064761P.
XX (BIOJ ) BIOGEN INC.
XX
XX Browning J;
XX WPI; 1999-418423/35.
XX N-PSDB; AAX59347.
XX
XX Novel B-cell myelin oligodendrocyte glycoproteins.
XX
XX Claim 2; Page 42; 43pp; English.
XX
XX This sequence represents human BMOG, a novel member of the B cell myelin
XX oligodendrocyte glycoprotein family that is expressed by germinal centre
XX B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were
XX identified. The protein is present primarily in the spleen, in lymph
XX nodes and in germinal centre B cells. It may have immunoregulatory
XX functions, and soluble or chimeric fusion proteins of BMOG may be used to
XX regulate the immune system in autoimmune or inflammatory disease. Vectors
XX comprising BMOG, prokaryotic and eukaryotic host cells, and a method of
XX producing BMOG using these transformed host cells are also provided. BMOG
XX polypeptides can be used for modulating the immune system of a subject or
XX to inhibit signal transduction in a cell expressing BMOG by contacting it
XX with a soluble BMOG protein. The nucleic acid can be used for gene
XX therapy. The protein can also be used to target a toxin, imaging agent or
XX radionuclide to a cell expressing BMOG. (All claimed)
XX
XX Sequence 190 AA;
XX
XX Query Match 99.8%; Score 1018; DB 2; Length 190;
XX Best Local Similarity 99.5%; Pred. No. 1e-91;
XX Matches 189; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MAWLLILLIMVPGSCALWVSPPEIRTLGSSAFPLPCSFNASQGRLAIGSVTFRDEV 60
DB 1 MAWLLILLIMVPGSCALWVSPPEIRTLGSSAFPLPCSFNASQGRLAIGSVTFRDEV 60
QY 61 VPGKEVNGTPEFRGRPLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
DB 61 VPGKEVNGTPEFRGRPLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120
QY 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYVYQKCHCHMGTHCHSSDGP 180
DB 121 NGTRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAVGSTVYVYQKCHCHMGTHCHSSDGP 180
QY 181 RGVIPEPRCP 190
DB 181 RGVIPEPRCP 190
XX
XX RESULT 3
XX AAY06403
XX ID AAY06403 standard; protein; 201 AA.

```



CC polypeptides can be used for modulating the immune system of a subject or  
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it  
 CC with a soluble BMOG protein. The nucleic acid can be used for gene  
 CC therapy. The protein can also be used to target a toxin, imaging agent or  
 CC radionuclide to a cell expressing BMOG. (All claimed)

XX SQ Sequence 177 AA;  
 Query Match 84.2%; Score 859; DB 2; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-76;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGLAIGSVTWRDEV 60  
 DB 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGLAIGSVTWRDEV 60  
 QY 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120  
 DB 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120  
 QY 121 NGTRLVVEKEHPOLGAGTVLLRAGEFYAVSFVSVAVGSTVYQGK 165  
 DB 121 NGTRLVVEKEHPOLGAGTVLLRAGEFYAVSFVSVAVGSTVYQGK 165

RESULT 5  
 AAE19109  
 ID AAE19109 standard; protein; 135 AA.  
 XX  
 AC AAE19109;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human NKp30 protein.  
 XX  
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;  
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;  
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; NKp30.  
 XX  
 OS Homo sapiens.  
 XX  
 OS Chimeric.  
 XX  
 FN WO200208287-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-IL000664.  
 XX  
 PR 20-JUL-2000; 2000IL-00137419.  
 XX  
 PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 PA (UYNE ) UNIV BEN-GURION NEGEV.  
 XX  
 PI Mandelboim O, Porzador A;  
 XX  
 XX WPI; 2002-195870/25.  
 DR N-PSDB; AAD30466.  
 XX  
 XX New targeting complex capable of targeting an active substance to a  
 PT target cell, comprising a target recognition segment and an active  
 PT segment, useful for treating pathologies associated with viral infections  
 PT or cancer.  
 XX  
 PS Example 1; Page 108; 113pp; English.

XX The invention relates to compositions and methods for the treatment and  
 CC detection of a variety of viral infections, by using complex agents  
 CC comprising the natural killer (NK) cells activating proteins, NKp46 and  
 CC NKp44 and functional fragments thereof, linked to therapeutic or imaging  
 CC agents. The complex is useful for treating pathologies associated with  
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein  
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECVV, MVM or herpes virus)  
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for  
 CC the imaging and monitoring of cancer. The complex may also be used to

CC detect the presence of abnormal cells in a sample. The antibodies can be  
 CC used to qualitatively or quantitatively detect the ligand for the  
 CC complex. The present sequence is human NKp30 protein

XX SQ Sequence 135 AA;  
 Query Match 69.9%; Score 713; DB 5; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-62;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGLAIGSVTWRDEV 60  
 DB 1 MAMWLLILIMVHPGSCALWVSOPPEIRITLGGSAFLPCSFNASQGLAIGSVTWRDEV 60  
 QY 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120  
 DB 61 VPGKEVRNGTPEFRGLAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGTG 120  
 QY 121 NGTRLVVEKEHPOLG 135  
 DB 121 NGTRLVVEKEHPOLG 135

RESULT 6  
 AAE19110  
 ID AAE19110 standard; protein; 369 AA.  
 XX  
 AC AAE19110;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human NKp30-IgG fusion protein.  
 XX  
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;  
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;  
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; IgG;  
 KW immunoglobulin G; fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Chimeric.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..135  
 FT /note= "Human NKp30"  
 FT Region 136..369  
 FT /note= "Human IgG"  
 XX  
 FN WO200208287-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-IL000664.  
 XX  
 PR 20-JUL-2000; 2000IL-00137419.  
 XX  
 PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 PA (UYNE ) UNIV BEN-GURION NEGEV.  
 XX  
 PI Mandelboim O, Porzador A;  
 XX  
 XX WPI; 2002-195870/25.  
 DR N-PSDB; AAD30467.  
 XX  
 XX New targeting complex capable of targeting an active substance to a  
 PT target cell, comprising a target recognition segment and an active  
 PT segment, useful for treating pathologies associated with viral infections  
 PT or cancer.  
 XX  
 PS Example 1; Page 108-110; 113pp; English.  
 XX The invention relates to compositions and methods for the treatment and  
 CC detection of a variety of viral infections, by using complex agents

comprising the natural killer (NK) cells activating proteins, Nkp46 and Nkp44 and functional fragments thereof, linked to therapeutic or imaging agents. The complex is useful for treating pathologies associated with viral infections (e.g. infections caused by influenza virus, HIV, Epstein-Barr virus, cytomegalovirus, vaccinia virus, ECV, MVM or herpes virus) and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for the imaging and monitoring of cancer. The complex may also be used to detect the presence of abnormal cells in a sample. The antibodies can be used to qualitatively or quantitatively detect the ligand for the NKp30 complex. The present sequence is human NKp30- immunoglobulin G (IgG) Fc region fusion protein. (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 369 AA;

Query Match 69.9%; Score 713; DB 5; Length 369;

Best Local Similarity 100.0%; Pred. No. 2.2e-61;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAWMLLLILIMVHPGSCALWVSOPPEIRTLGSSAFPCSFNASQGRLAIGSVTFRDEV 60

DB 1 MAWMLLLILIMVHPGSCALWVSOPPEIRTLGSSAFPCSFNASQGRLAIGSVTFRDEV 60

QY 61 VPGKEVRNGTPEFRGRFLAPLASSRFLHDHQAELHIRDVRGHDASYVCRVEVLGLGVGTG 120

DB 61 VPGKEVRNGTPEFRGRFLAPLASSRFLHDHQAELHIRDVRGHDASYVCRVEVLGLGVGTG 120

QY 121 NGTRLVVEKEHPOLG 135

DB 121 NGTRLVVEKEHPOLG 135

RESULT 7

AAE02771

ID AAE02771 standard; protein; 120 AA.

XX

AC AAE02771;

XX

DT 06-AUG-2001 (first entry)

XX

DE Human NKp30 receptor extracellular region sequence.

XX

KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;

KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

KW therapy; extracellular region.

XX

OS Homo sapiens.

XX

FN WO200136630-A2.

XX

PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-EP011697.

XX

PR 15-NOV-1999; 99CA-02288307.

XX

PR 15-NOV-1999; 99US-00440514.

XX

PA (INNA-) INNATE PHARMA SAS.

XX

PA (UYGE-) UNIV GENOVA.

XX

PI Moretta A, Bottino C, Biassoni R;

XX

DR WPI; 2001-329221/34.

XX

PT Novel compound, useful for detection and/or quantifying the presence of

XX NK cells, comprises the amino acid sequences of the NKp30 molecule.

XX

PS Claim 1; Fig 7B; 83pp; English.

XX

CC The invention relates to human NKp30 receptor and its corresponding cDNA

XX molecule which is involved in natural cytotoxicity mediated by natural

CC killer (NK) cells and antibodies that identify the same. NKp30 receptor

XX is a member of immunoglobulin super family (Ig-SF). NKp30 is selectively

expressed on the surface of human mature NK cells. NKp30 and its cDNA are useful for detecting and/or quantifying the presence of NK cells in a biological sample. The invention also provides kits for detecting and/or quantifying the presence of NK cells, for the selective removal of NK cells from a biological sample, for the positive and selective purification of NK cells from a biological sample and for the in vitro stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (GVT) and especially graft versus leukaemia (GVL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. NKp30 antibodies are useful for identifying NKp30 natural ligands and allow assessment of the level of surface NKp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence NKp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the extracellular region of human NKp30 receptor

XX Sequence 120 AA;

Query Match 62.0%; Score 632; DB 4; Length 120;

Best Local Similarity 100.0%; Pred. No. 4.7e-54;

Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 LWSVOPPEIRTLGSSAFPCSFNASQGRLAIGSVTFRDEVVPGKEVRNGTPEFRGLA 78

DB 1 LWSVOPPEIRTLGSSAFPCSFNASQGRLAIGSVTFRDEVVPGKEVRNGTPEFRGLA 60

QY 79 PLASSRFLHDHQAELHIRDVRGHDASYVCRVEVLGLGVGTGNTGLVVEKEHPOLGAGT 138

DB 61 PLASSRFLHDHQAELHIRDVRGHDASYVCRVEVLGLGVGTGNTGLVVEKEHPOLGAGT 120

RESULT 8

AAE02773

ID AAE02773 standard; peptide; 33 AA.

XX

AC AAE02773;

XX

DT 06-AUG-2001 (first entry)

XX

DE Human NKp30 receptor intracellular region sequence.

XX

KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;

KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

KW therapy; intracellular region.

XX

OS Homo sapiens.

XX

FN WO200136630-A2.

XX

PD 25-MAY-2001.

XX

PF 15-NOV-2000; 2000WO-EP011697.

XX

PR 15-NOV-1999; 99CA-02288307.

XX

PR 15-NOV-1999; 99US-00440514.

XX

PA (INNA-) INNATE PHARMA SAS.

XX

PA (UYGE-) UNIV GENOVA.

XX

PI Moretta A, Bottino C, Biassoni R;

XX

DR WPI; 2001-329221/34.

XX

PT Novel compound, useful for detection and/or quantifying the presence of

XX NK cells, comprises the amino acid sequences of the NKp30 molecule.

XX

PS Claim 1; Fig 7B; 83pp; English.

XX

CC The invention relates to human NKp30 receptor and its corresponding cDNA



CC molecule which is involved in natural cytotoxicity mediated by natural  
 CC killer (NK) cells and antibodies that identify the same. NKP30 receptor  
 CC is a member of immunoglobulin super family (Ig-SF). NKP30 is selectively  
 CC expressed on the surface of human mature NK cells. NKP30 and its cDNA are  
 CC useful for detecting and/or quantifying the presence of NK cells in a  
 CC biological sample. The invention also provides kits for detecting and/or  
 CC quantifying the presence of NK cells, for the selective removal of NK  
 CC cells from a biological sample, for the positive and selective  
 CC purification of NK cells from a biological sample and for the in vitro  
 CC stimulation of NK cell cytotoxicity. The invention further provides a  
 CC pharmaceutical composition which is used as a drug for grafting  
 CC enhancement, graft versus host (GvH) inhibition, stimulation of graft  
 CC versus tumour (GvT) and especially graft versus leukaemia (GvL), and for  
 CC the prevention, palliation and/or therapy of solid or liquid tumours,  
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or  
 CC microorganism, notably viral infection. NKP30 antibodies are useful for  
 CC identifying NKP30 natural ligands and allow assessment of the level of  
 CC surface NKP30 ligand expressed on an NK-susceptible target cell and the  
 CC comparison of this level to the standard physiological one. Hence NKP30  
 CC antibodies are useful in the diagnosis of tumours or of infection. The  
 CC present sequence is the intracellular region of human NKP30 receptor  
 XX  
 SQ Sequence 33 AA;

Query Match 20.0%; Score 204; DB 4; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-13;  
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 STVYQKCHCHGTHCHSSDGRGVIPRCP 190  
 DB 1 STVYQKCHCHGTHCHSSDGRGVIPRCP 33

RESULT 9  
 AAE20273  
 ID AAE20273 standard; protein; 246 AA.  
 AC AAE20273;  
 XX  
 XX 18-JUN-2002 (first entry)  
 DT  
 DE Human lung specific gene (LSG) protein #11.  
 XX  
 XX Human; lung specific gene; LSG; lung embryonic development; cytostatic;  
 KW lung cancer; vaccine; gene therapy; non-cancerous lung disease; tumour.  
 KW  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 PH Region 189..204  
 FT /note= "Antigenic epitope"  
 FT  
 XX WO200208278-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 20-JUL-2001; 2001WO-US022949.  
 XX  
 XX 21-JUL-2000; 2000US-0219834P.  
 XX  
 XX (DIAD-) DIADEXUS INC.  
 XX  
 XX Macina RA, Nair M, Chen S;  
 XX  
 XX WPI; 2002-268964/31.  
 XX  
 XX Novel lung specific gene useful for identifying, diagnosing, monitoring,  
 PT staging, imaging and treating lung cancer and non-cancerous disease  
 PT states in lung, for gene therapy, and for identifying lung tissue.  
 PT  
 XX Claim 2; Page 178-179; 197pp; English.  
 PS  
 XX The present invention relates to lung specific genes (LSG) and their

CC corresponding polypeptides. LSG is useful for identifying, diagnosing,  
 CC monitoring, staging, imaging and treating lung cancer and non-cancerous  
 CC disease states in lung, identifying lung tissue, monitoring and modifying  
 CC lung embryonic development and differentiation, in gene therapy, as  
 CC hybridisation probes, to detect LSG mRNA as a marker for lung cancer, as  
 CC research reagents and materials for discovery of treatments and  
 CC diagnostics to human disease, to detect complementary polynucleotides,  
 CC and for chromosome identification. An antibody which binds LSG is useful  
 CC to detect or image localisation of LSG in a patient for detecting or  
 CC diagnosing a disease or condition, for preventing the onset and treatment  
 CC of lung cancer, to isolate or to identify clones expressing LSG  
 CC polypeptides, to purify LSG polypeptides, and to target tumours  
 CC expressing LSG. The present sequence is human LSG protein  
 XX  
 SQ Sequence 246 AA;

Query Match 12.1%; Score 123.5; DB 5; Length 246;  
 Best Local Similarity 29.2%; Pred. No. 0.001;  
 Matches 54; Conservative 17; Mismatches 61; Indels 53; Gaps 10;

QY 1 MAMM-LILLIMVHFGSCALWV-SQPETITLGSAGFLPCSNASQGRLAIGSVTWFRD 58  
 DB 1 MAMAPLLLTLLSLTGLSLQFILTQPPSASASLGASVTLTCSVSSDYKNL---EVDWFOQ 57

QY 59 EVVPGKEVR-----NGTPEFRGLAP-----LASSRFLHDHQAELHIRDVRGHDA 104  
 DB 58 R--PKGPRFVMRVGTGVVGFGRGADIPDRFSVSGSLARF-----LTIRIEDEDS 108

QY 105 IYCVREVLGLGVGT-----NGTSLVV-----EKEHPQLGAGTVLL 142  
 DB 109 DYHCGTD---LGSSTFSVSWVFGGTLTIVLSQPKAAPSVTLPFPSPSEELQANKATIVCL 165

QY 143 RAGFY 147  
 DB 166 ISDFY 170

RESULT 10  
 ABU08020  
 ID ABU08020 standard; protein; 232 AA.  
 AC ABU08020;  
 XX  
 XX 10-MAY-2003 (first entry)  
 DT  
 DE Monoclonal rabies virus antibody light chain, clone JB.1.  
 XX  
 KW Human; antibody; constant region; monoclonal antibody 57; Mab 57;  
 KW variable region; Rabies; neurological disease; infection;  
 KW central nervous system; rabies virus; Lyssavirus; Rhabdoviridae;  
 KW pathogen; vaccine; virucide; light chain.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO2003016501-A2.  
 XX  
 XX 27-FEB-2003.  
 XX  
 XX 21-AUG-2002; 2002WO-US026584.  
 XX  
 XX 21-AUG-2001; 2001US-0314023P.  
 XX  
 XX (UYJE-) UNIV JEFFERSON THOMAS.  
 XX  
 XX Hooper DC, Dietzschold B;  
 XX  
 XX WPI; 2003-278566/27.  
 DR N-PSDB; ABX12862.  
 XX  
 XX New recombinant antibody comprising a constant region of Mab 57 linked to  
 PT a non-Mab 57 variable region, useful for treating an individual exposed  
 PT to a pathogen, e.g. rabies infection.  
 XX



```

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PA
XX Linsley PS, Ledbetter JA, Damle NK, Brady W, Wallace PM;
XX PI
XX WPI; 1994-272160/34.
XX DR N-PSDB; AAQ70451.
XX DR
XX PT Use of a CTLA4-binding molecule and an IL4-binding molecule - for
XX PT regulating an immune response involving B7-positive lymphocytes, partic.
XX PT in transplant rejection.
XX PT
XX Example 3; Fig 3; 96pp; English.
XX PS
XX This sequence shows the human CTLA4 receptor. The cDNA encoding it was
XX CC assembled from 2 PCR fragments amplified from mRNA from H38 leukaemia
XX CC cells. The CTLA4 receptor is express using the the oncostatin M signal
XX CC peptide fused to the N terminus of CTLA4. The methods of the invention
XX CC can be used for regulating a B cell response resulting in the inhibition
XX CC of antibody produ., regulating a T cell response resulting in the
XX CC inhibition of cell mediated immunity or for the inhibition of lymphocyte
XX CC proliferation (claimed)
XX CC
XX Sequence 212 AA;
XX SQ
XX Query Match 11.6%; Score 118.5; DB 2; Length 212;
XX Best Local Similarity 28.1%; Pred. No. 0.0026;
XX Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;
XX
XX QY 4 MLLLLILIMVHP--GSCALWVSQPPEIRTLGSSAFLPCSFNASQGRLAIGSVTFWRDEW 61
XX Db 10 LLSLLVALLFFTSMAAMHVAQPAVLASSRGIAFVCEY-ASPGKATEVRVTVLRLQADS 68
XX QY 62 PGKEVRNQTPEFRGELAPLASSRLHP-----HQAEHLTRDVRGHDSIYVCREV 112
XX Db 69 QVTEVCAATYMMGNELT-----FLDPSICTGSSGNQNLITQGLRAMDTGLYICKVEL 122
XX QY 113 L---GLGVGTGNGTRLVYVEKEHPQLGAGTVLLLRA----GFYAVSPISVAV 156
XX Db 123 MYPPPPYVLGIGNGTQIVVIDPEPCPSDFLLWILAAVSSGLPFYSFLTAV 173
XX
XX RESULT 13
XX AAR77642
XX ID AAR77642 standard; protein; 212 AA.
XX AC
XX AC AAR77642;
XX DT
XX DT 11-JUN-1996 (first entry)
XX DE
XX DE Full length CTLA4 and oncostatin M signal peptide.
XX KW primer; PCR; polymerase chain reaction; CTLA4; CD28; B7 antigen;
XX KW fusion protein; regulation; immune response; lymphocyte; oncostatin M;
XX KW soluble; chimeric; inhibit; transplant rejection;
XX KW graft versus host disease.
XX XX
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH 1..25
XX FT /label= oncostatin_M_signal_peptide
XX FT 26
XX FT /label= CTLA4
XX FT Modified-site 134..136
XX FT /note= "glycosylation site"
XX XX
XX DN AU9516458-A.
XX XX
XX XX 26-OCT-1995.
XX PD
XX PD 13-APR-1995; 95AU-00016458.
XX PF
XX XX

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 66.0212 Seconds  
(without alignments)  
607.670 Million cell updates/sec

Title: US-10-036-444-2  
Perfect score: 1020  
Sequence: 1 MAWMLLLILMVHGPSALW.....GTHCHSSDGRGVIPEPRCP 190

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUBCOMB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	190	13	US-10-036-444-2
2	632	62.0	120	13	US-10-036-444-4
3	204	20.0	33	13	US-10-036-444-6
4	123.5	12.1	246	10	US-09-909-567B-49
5	119.5	11.7	232	14	US-10-225-108A-12
6	119.5	11.7	232	15	US-10-461-148-6
7	118.5	11.6	139	14	US-10-312-495-6
8	118.5	11.6	212	10	US-09-898-195A-17
9	118.5	11.6	212	14	US-10-057-288-12
10	118.5	11.6	212	14	US-10-155-514-2
11	118.5	11.6	212	16	US-10-419-008-17
12	111.5	10.9	226	13	US-10-038-107A-1
13	111.5	10.9	226	15	US-10-371-069-32
14	111.5	10.9	226	15	US-10-371-645-32
15	111.5	10.9	226	15	US-10-371-260-32

16	111.5	10.9	226	15	US-10-411-010-19	Sequence 19, Appl
17	110	10.8	223	9	US-09-989-545-21	Sequence 21, Appl
18	110	10.8	223	10	US-09-928-267-18	Sequence 18, Appl
19	110	10.8	223	10	US-09-928-267-21	Sequence 21, Appl
20	110	10.8	223	14	US-10-225-519-8	Sequence 8, Appl
21	110	10.8	223	14	US-10-207-655-101	Sequence 101, App
22	108.5	10.6	223	9	US-09-989-545-20	Sequence 20, Appl
23	108.5	10.6	223	14	US-10-211-207-5	Sequence 5, Appl
24	108.5	10.6	223	14	US-10-077-106-5	Sequence 5, Appl
25	108.5	10.6	223	10	US-09-898-195A-11	Sequence 11, Appl
26	108.5	10.6	383	14	US-10-057-288-6	Sequence 6, Appl
27	108.5	10.6	383	14	US-10-155-514-10	Sequence 10, Appl
28	108.5	10.6	383	16	US-10-419-008-11	Sequence 11, Appl
29	108	10.6	223	13	US-10-107-828-26	Sequence 26, Appl
30	108	10.6	223	13	US-10-107-907-26	Sequence 26, Appl
31	108	10.6	223	13	US-10-107-868-26	Sequence 26, Appl
32	108	10.6	223	14	US-10-301-056-26	Sequence 26, Appl
33	108	10.6	260	10	US-09-928-267-13	Sequence 13, Appl
34	108	10.6	260	10	US-09-928-267-14	Sequence 14, Appl
35	108	10.6	260	14	US-10-225-519-4	Sequence 4, Appl
36	107.5	10.5	261	10	US-09-928-267-9	Sequence 9, Appl
37	107.5	10.5	261	10	US-09-928-267-10	Sequence 10, Appl
38	107.5	10.5	261	14	US-10-225-519-2	Sequence 2, Appl
39	107	10.5	223	10	US-09-835-297-2	Sequence 2, Appl
40	107	10.5	223	14	US-10-211-207-3	Sequence 3, Appl
41	107	10.5	223	14	US-10-077-106-3	Sequence 3, Appl
42	107	10.5	223	15	US-10-336-384-7	Sequence 7, Appl
43	107	10.5	382	14	US-10-207-655-307	Sequence 307, App
44	106.5	10.4	151	14	US-10-207-655-309	Sequence 309, App
45	106.5	10.4	223	10	US-09-928-267-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-10-036-444-2  
; Sequence 2, Application US/10036444  
; Publication No. US2002014245A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; APPLICANT: UNIVERSITA DI GENOVA  
; TITLE OF INVENTION: "No. US2002014245A1 triggering receptor involved in natural  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and  
; TITLE OF INVENTION: antibodies that identify the same"  
; FILE REFERENCE: SEQ-FR-1060  
; CURRENT APPLICATION NUMBER: US/10/036,444  
; PRIOR FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: 09/440,514  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: 09/456,199  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Human NK cell  
US-10-036-444-2

Query Match	100.0%;	Score 1020;	DB 13;	Length 190;
Best Local Similarity	100.0%;	Pred. No. 2e-94;		
Matches 190;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAWMLLLILMVHGPSALWVSQPPFTELGSSAFUPCSFNASQGLAIGSVTWPFDEV	60	
Db	1	MAWMLLLILMVHGPSALWVSQPPFTELGSSAFUPCSFNASQGLAIGSVTWPFDEV	60	
QY	61	VPGEVNGTPEFRGRAPLASSRFLHDHQAELHVRGHDASIVYCRVENVGLGVGTG	120	
Db	61	VPGEVNGTPEFRGRAPLASSRFLHDHQAELHVRGHDASIVYCRVENVGLGVGTG	120	
QY	121	NGTRLVVEKEHPQLGAGTVLLRAGFYAVSFVSVVGYQKCHCHNGTHCHSSDGP	180	

Db 121 NGRLVVEKEHPQLGAGTVLLLRAGFYAVSFLSVAGSVTYVYQKCHCHGTHCHSSDGP 180  
QY 181 RGVIPERCP 190  
Db 181 RGVIPERCP 190

RESULT 2  
US-10-036-444-4  
; Sequence 4, Application US/10036444  
; Publication No. US20020142445A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; APPLICANT: UNIVERSITA DI GENOVA  
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and  
; TITLE OF INVENTION: antibodies that identify the same"  
; FILE REFERENCE: SEQ-PR-1060  
; CURRENT APPLICATION NUMBER: US/10/036,444  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: 09/440,514  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: 09/456,199  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patentin ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Human NK cell  
US-10-036-444-4

Query Match 52.0%; Score 632; DB 13; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.1e-55;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 19 LWSQPEIRTLGSSAFPCSFNASQGLAIGSVTWPRDEVVPGKEVRNGTPEERGLA 78  
Db 1 LWSQPEIRTLGSSAFPCSFNASQGLAIGSVTWPRDEVVPGKEVRNGTPEERGLA 60  
QY 79 PLASSRPLHDHQAELHIRDVRGHDAIYICRVEVLGLVGTGNGRLVVEKEHPQLGAGT 138  
Db 61 PLASSRPLHDHQAELHIRDVRGHDAIYICRVEVLGLVGTGNGRLVVEKEHPQLGAGT 120

RESULT 3  
US-10-036-444-6  
; Sequence 6, Application US/10036444  
; Publication No. US20020142445A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; APPLICANT: UNIVERSITA DI GENOVA  
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and  
; TITLE OF INVENTION: antibodies that identify the same"  
; FILE REFERENCE: SEQ-PR-1060  
; CURRENT APPLICATION NUMBER: US/10/036,444  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: 09/440,514  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: 09/456,199  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human NK cell  
US-10-036-444-6  
Query Match 20.0%; Score 204; DB 13; Length 33;  
Best Local Similarity 100.0%; Pred. No. 2e-13;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 158 STVYQKCHCHGTHCHSSDGRGVIPERCP 190  
Db 1 STVYQKCHCHGTHCHSSDGRGVIPERCP 33

RESULT 4  
US-09-909-567B-49  
; Sequence 49, Application US/09909567B  
; Publication No. US2003002257A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto A.  
; APPLICANT: Nair, Manoj  
; APPLICANT: Chen, Selyu  
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes  
; FILE REFERENCE: DX-0214  
; CURRENT APPLICATION NUMBER: US/09/909,567B  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: 60/219,834  
; PRIOR FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 49  
; LENGTH: 246  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-909-567B-49

Query Match 12.1%; Score 123.5; DB 10; Length 246;  
Best Local Similarity 29.2%; Pred. No. 0.00031;  
Matches 54; Conservative 17; Mismatches 61; Indels 53; Gaps 10;  
QY 1 MAMV-LLLILIMVHPCSCALWV-SQPEIRTLGSSAFPCSFNASQGLAIGSVTWFRD 58  
Db 1 MAMVPLLLTLLSULTGSLSPILTPPSASASLGASVTLTCSVSSDYKYL---EVDWFOQ 57  
QY 59 EVVPGKEVR-----NGTPEERGLAP-----LASSRPLHDHQAELHIRDVRGHDA 104  
Db 58 R--PGKPRFVVRVGTGVVGFGRGADIPDRFSVSGSLNRF-----LTIRNIEEDES 108  
QY 105 IYVCRVRLGLGVGT-----GNGTRLVY-----EKEHPQLGAGTVLL 142  
Db 109 DYHCGTD--LGSSTSVSVWVFGGKTLTVLSQPKAAPSVTLPFPSPSEELQANKATIVCL 165  
QY 143 RAGFY 147  
Db 166 ISDFY 170

RESULT 5  
US-10-225-108A-12  
; Sequence 12, Application US/10225108A  
; Publication No. US20030157112A1  
; GENERAL INFORMATION:  
; APPLICANT: HOOPER, Craig  
; APPLICANT: DIETZSCHOLD, Bernhard  
; TITLE OF INVENTION: Recombinant Antibodies, and Compositions  
; TITLE OF INVENTION: and Methods for Making Them  
; FILE REFERENCE: 8321-110  
; CURRENT APPLICATION NUMBER: US/10/225,108A  
; CURRENT FILING DATE: 2003-04-10  
; PRIOR APPLICATION NUMBER: US 09/848,832  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/204,518  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: US 60/314,023  
; PRIOR FILING DATE: 2001-08-21  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 232  
; TYPE: PRT

```
; ORGANISM: Homo sapiens
US-10-225-108A-12

Query Match
  11.7%; Score 119.5; DB 14; Length 232;
Best Local Similarity 26.0%; Pred. No. 0.00074;
Matches 45; Conservative 25; Mismatches 64; Indels 39; Gaps 7;

QY 1 MAMMLLLIMVH-PGSCALWV-SQPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRD 58
Db 1 MAWTVLLGLSHCTGSVTSYLTQPPSVAPGKTARINCNNIEYR-----SVHWYQQ 56

QY 59 E-----VPEKVRNGTPE-FRGRAPLASSRFLHDHQAELHIRDVRGHDASIYVCR 109
Db 57 KSGQAPVAVIYDNRPSGIPERFSGSKS-----GNTATLTISRVEAGDEADYTCQ 107

QY 110 VEVLGLGVTGNGTRLVW-----EKEHPOLGAGTVLLLRAGPY 147
Db 108 VNDISSDVVFGGTKLTVLGQKAPSVTLFPPSSEELQANKATIVCLISDFY 160

RESULT 6
US-10-461-148-6
; Sequence 6, Application US/10461148
; Publication No. US20040013672A1
; GENERAL INFORMATION:
; APPLICANT: Hooper, Douglas C.
; TITLE OF INVENTION: RECOMBINANT ANTIBODIES AND COMPOSITIONS
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THE SAME
; FILE REFERENCE: 8321-110C11-185685
; CURRENT APPLICATION NUMBER: US/10/461,148
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 10/225,108
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: US 60/314,023
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/848,832
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/204,518
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; TYPE: PRT
; ORGANISM: Human
US-10-461-148-6

Query Match
  11.7%; Score 119.5; DB 15; Length 232;
Best Local Similarity 26.0%; Pred. No. 0.00074;
Matches 45; Conservative 25; Mismatches 64; Indels 39; Gaps 7;

QY 1 MAMMLLLIMVH-PGSCALWV-SQPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRD 58
Db 1 MAWTVLLGLSHCTGSVTSYLTQPPSVAPGKTARINCNNIEYR-----SVHWYQQ 56

QY 59 E-----VPEKVRNGTPE-FRGRAPLASSRFLHDHQAELHIRDVRGHDASIYVCR 109
Db 57 KSGQAPVAVIYDNRPSGIPERFSGSKS-----GNTATLTISRVEAGDEADYTCQ 107

QY 110 VEVLGLGVTGNGTRLVW-----EKEHPOLGAGTVLLLRAGPY 147
Db 108 VNDISSDVVFGGTKLTVLGQKAPSVTLFPPSSEELQANKATIVCLISDFY 160

RESULT 7
US-10-312-495-6
; Sequence 6, Application US/10312495
; Publication No. US20030165495A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: A0999PCT000454-124

; CURRENT APPLICATION NUMBER: US/10/312,495
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: 60/213,630
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/283,813
; PRIOR FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-312-495-6

Query Match
  11.6%; Score 118.5; DB 14; Length 139;
Best Local Similarity 30.2%; Pred. No. 0.00048;
Matches 39; Conservative 17; Mismatches 40; Indels 33; Gaps 7;

QY 20 WVS-----QPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDE-----VWPG 63
Db 18 WWSGDQVKQSPSALSLEGTSSALRCNFS-----IATTVQNFQNSRGLMNLFLVLP- 71

QY 64 KEVRNGTPEFRGRAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVVLGLG--VGTGN 121
Db 72 -----GTKE-NGRLKSTFNSK---ESYSTLHIRDAQLEDSTGYFCAARVEGTGSKLSFGK 122

QY 122 GTRLVVEKE 130
Db 123 GAKLTWSPD 131

RESULT 8
US-09-898-195A-17
; Sequence 17, Application US/09898195A
; Publication No. US20030083246A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Robert
; APPLICANT: Carr, Suzette
; APPLICANT: Hagerty, David
; APPLICANT: Beach, Robert J
; APPLICANT: Becker, Jean-Claude
; TITLE OF INVENTION: METHODS FOR TREATING RHEUMATIC DISEASES USING A SOLUBLE
; TITLE OF INVENTION: CTLA4 MOLECULE
; FILE REFERENCE: D0030NP/30436.55USU1
; CURRENT APPLICATION NUMBER: US/09/898,195A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/215,913
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-898-195A-17

Query Match
  11.6%; Score 118.5; DB 10; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLLLIMVHP--GSCALWVSQPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVV 61
Db 10 LLSLVLLALLFPSSMASMAHVAQPAVVLASSRGIAFSVCEY-ASPGKATEVRYTVLQRADS 68

QY 62 PGKEVRNGTPEFRGRAPLASSRFLHD-----HQAEHLIRDVRGHDASIYVCRV 112
Db 69 QVTEVCAATYMMGNELT-----FLDDSTCTGTSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLLR-----GFYAVSFLSVAV 156
Db 123 MYPPPYVLGNGTQIYVIDPECPDSDFLMLILAVSSGLFFYFLLTAV 173
```



```
RESULT 9
US-10-057-288-12
; Sequence 12, Application US/10057288
; Publication No. US20030007968A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Christian P.
; APPLICANT: Pearson, Thomas C.
; APPLICANT: Waller, Edmund K.
; APPLICANT: Adams, Andrew B.
; TITLE OF INVENTION: METHODS OF INDUCING ORGAN TRANSPLANT TOLERANCE AND
; FILE REFERENCE: D0136NP/30436.58USUI
; CURRENT APPLICATION NUMBER: US/10/057,288
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/264,528
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/303,142
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 12
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-288-12

Query Match 11.6%; Score 118.5; DB 14; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLILIMVHP--GSCALWUSQPEIRTELEGSSAFPCSNASQGRLAIGSVTWFRDEVV 61
DB 10 LLSLVIALFFPSMASMAHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLQADS 58

QY 62 PGKEVRNGTPEFRGRGLAPLASSRFLHD-----HQAELHIRDVRGHDASIYVCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA---CFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173

US-10-155-514-2
; Sequence 2, Application US/10155514
; Publication No. US20030022836A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Christian P.
; APPLICANT: Pearson, Thomas C.
; APPLICANT: Adams, Andrew B.
; TITLE OF INVENTION: METHODS FOR PROTECTING ALLOGENEIC ISLET TRANSPLANT USING SOLUBLE
; FILE REFERENCE: D0173NP / 30436.62USUI
; CURRENT APPLICATION NUMBER: US/10/155,514
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 60/293,402
; PRIOR FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-155-514-2

Query Match 11.6%; Score 118.5; DB 14; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLILIMVHP--GSCALWUSQPEIRTELEGSSAFPCSNASQGRLAIGSVTWFRDEVV 61
DB 10 LLSLVIALFFPSMASMAHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLQADS 58

QY 62 PGKEVRNGTPEFRGRGLAPLASSRFLHD-----HQAELHIRDVRGHDASIYVCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA---CFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173
```

```
QY 62 PGKEVRNGTPEFRGRGLAPLASSRFLHD-----HQAELHIRDVRGHDASIYVCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA---CFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173

RESULT 11
US-10-419-008-17
; Sequence 17, Application US/10419008
; Publication No. US20040022787A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Robert
; APPLICANT: Carr, Suzette
; APPLICANT: Hagerty, David
; APPLICANT: Peach, Robert J.
; APPLICANT: Becker, Jean-Claude
; TITLE OF INVENTION: METHODS FOR TREATING AN AUTOIMMUNE DISEASE USING A SOLUBLE CTLAA
; FILE REFERENCE: 30436.55USUI
; CURRENT APPLICATION NUMBER: US/10/419,008
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: 09/896,195
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 60/215,913
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/373,852
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/407,246
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-419-008-17

Query Match 11.6%; Score 118.5; DB 16; Length 212;
Best Local Similarity 28.1%; Pred. No. 0.00083;
Matches 48; Conservative 24; Mismatches 74; Indels 25; Gaps 6;

QY 4 MLLILIMVHP--GSCALWUSQPEIRTELEGSSAFPCSNASQGRLAIGSVTWFRDEVV 61
DB 10 LLSLVIALFFPSMASMAHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLQADS 68

QY 62 PGKEVRNGTPEFRGRGLAPLASSRFLHD-----HQAELHIRDVRGHDASIYVCRVEV 112
DB 69 QVTEVCAATYMMGNELT-----FLDDSICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122

QY 113 L---GLGVGTGNGTRLVVEKEHPOLGAGTVLLRA---CFYAVSFSLVAV 156
DB 123 MYPPPYLIGNGTQIYVIDPEPCDSDFLWLILAAVSSGLFFYSFLLTAV 173

RESULT 12
US-10-038-107A-1
; Sequence 1, Application US/10038107A
; Publication No. US20020150573A1
; GENERAL INFORMATION:
; APPLICANT: Nussenzweig, Michel
; TITLE OF INVENTION: ANTI-1G ALPHA-BETA ANTIBODY FOR LYMPHOMA THERAPY
; FILE REFERENCE: 7529/OH405
; CURRENT APPLICATION NUMBER: US/10/038,107A
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/247,079
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
```

```
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-107A-1

Query Match      10.9%; Score 111.5; DB 13; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILLIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
Db 18 LFLLSAVYLGPGCOALWMHKVPASLVSLGDEAHFQCPHNSNN-----ANVTWR--VLH 71
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLIRDVGRGHASIIYVCRVEVLGLGVGTG 120
Db 72 G-----NYTWPEFLGP-----GEDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
Db 119 CGTYLRVQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLILFR 166

RESULT 13
US-10-371-069-32
; Sequence 32, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,069
; CURRENT FILING DATE: 2003-02-21
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-069-32

Query Match      10.9%; Score 111.5; DB 15; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILLIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
Db 18 LFLLSAVYLGPGCOALWMHKVPASLVSLGDEAHFQCPHNSNN-----ANVTWR--VLH 71
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLIRDVGRGHASIIYVCRVEVLGLGVGTG 120
Db 72 G-----NYTWPEFLGP-----GEDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
Db 119 CGTYLRVQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLILFR 166

RESULT 14
US-10-371-645-32
; Sequence 32, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371,645
; CURRENT FILING DATE: 2003-06-20
; PRIOR FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-645-32

Query Match      10.9%; Score 111.5; DB 15; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILLIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
Db 18 LFLLSAVYLGPGCOALWMHKVPASLVSLGDEAHFQCPHNSNN-----ANVTWR--VLH 71
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLIRDVGRGHASIIYVCRVEVLGLGVGTG 120
Db 72 G-----NYTWPEFLGP-----GEDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
Db 119 CGTYLRVQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLILFR 166

RESULT 15
US-10-371-260-32
; Sequence 32, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371,260
; CURRENT FILING DATE: 2003-02-21
; PRIOR FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human MB-1 Ig-alpha
US-10-371-069-32

Query Match      10.9%; Score 111.5; DB 15; Length 226;
Best Local Similarity 28.0%; Pred. No. 0.0045;
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;

QY 4 MLLILLIMVHFGSCALWVSQ--PPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVP 62
Db 18 LFLLSAVYLGPGCOALWMHKVPASLVSLGDEAHFQCPHNSNN-----ANVTWR--VLH 71
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAEHLIRDVGRGHASIIYVCRVEVLGLGVGTG 120
Db 72 G-----NYTWPEFLGP-----GEDPNGTLIIQNVKSHGGIYVCRVQEGNESYQOS 118
QY 121 NGTRLVVEKEHPQ-----LGAGTV--LLLRAGFYAVSFLSVAGSVTVYQ 163
Db 119 CGTYLRVQPPRPFLDMGEGTKNRIITAGIILLFCVAVPGTLLILFR 166
```





Query Match 11.0%; Score 112; DB 3; Length 223;  
Best Local Similarity 28.0%; Pred. No. 0.00046;  
Matches 47; Conservative 22; Mismatches 75; Indels 24; Gaps 6;  
QY 6 LLLILIMVHSGC-ALWVSQPEIRTLRGSSAFPCSFNASQGLRAIGSVTWFRDEVPGK 64  
DB 24 LLLFLLFIPVCKAMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLRQADSQVT 82  
QY 65 EVRNGTPEFRGLAPLASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL-- 113  
DB 83 EVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYP 136  
QY 114 -GLGVGTGNGTRLVVEKHPQLGAGTVLLRA-----GFYAVSFLSVAV 156  
DB 137 PPTVIGLGNQTIYVIDPEPCDFDILLWILAAVSSGLFFYSFLLTAV 184

RESULT 2  
US-09-311-784A-32  
; Sequence 32, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-2022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 32  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human MB-1 Ig-alpha  
US-09-311-784A-32

Query Match 10.9%; Score 111.5; DB 4; Length 226;  
Best Local Similarity 28.0%; Pred. No. 0.00053;  
Matches 47; Conservative 24; Mismatches 70; Indels 27; Gaps 8;  
QY 4 MLLILIMVHSGCALWVSQ-PPEIRTEGSSAFPCSFNASQGLRAIGSVTWFRDEVYP 62  
DB 18 LFLLSAVYLGPGCOALMWHKVPASLVSLGEDAHFQCPHNSNN---ANVTWVR--VLH 71  
QY 63 GKEVRNGT--PEFRGLAPLASSRFLHDQAELHIRDVRGHDSIYVCRVEVLGLGVGTG 120  
DB 72 G-----NTVPEFLGP-----GEDPNTGTLIQNVKSHGGIYVCRVQEGNESYQOS 118  
QY 121 NGRVLVVEKHPQ-----LGAGTV-LLLRAGFYAVSFLSVAGSVTVYQ 163  
DB 119 CGTVLRVRQPPRPFLDMGSGTKNRIITASGIILLFCVAVPGTLLFLFR 166

RESULT 3  
US-09-472-087-100  
; Sequence 100, Application US/09472087  
; Patent No. 6682736  
; GENERAL INFORMATION:  
; APPLICANT: HANSON, DOUGLAS C.  
; APPLICANT: NEVEU, MARK J.  
; APPLICANT: MUELLER, EILEEN E.  
; APPLICANT: HANKE, JEFFREY H.

; APPLICANT: GILMAN, STEVEN C.  
; APPLICANT: DAVIS, C. GEOFFREY  
; APPLICANT: CORVALAN, JOSE R.  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4  
; FILE REFERENCE: ABX-PFI  
; CURRENT APPLICATION NUMBER: US/09/472,087  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/113,647  
; PRIOR FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 147  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 100  
; LENGTH: 364  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-472-087-100

Query Match 10.4%; Score 106.5; DB 4; Length 364;  
Best Local Similarity 28.0%; Pred. No. 0.0035;  
Matches 40; Conservative 20; Mismatches 62; Indels 21; Gaps 5;  
QY 4 MLLILIMVHP--GSCALWVSQPEIRTEGSSAFPCSFNASQGLRAIGSVTWFRDEVV 61  
DB 10 LLSLVIALLLFFPSMASMAHVAQPAVVLASSRGIAFVCEY-ASPGKATEVRVTVLRQADS 68  
QY 62 PKQVRNGTPEFRGLAPLASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEV 112  
DB 69 QVTEVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVEL 122  
QY 113 L---GLGVGTGNGTRLVVEKEHP 132  
DB 123 MYPPPYLIGNGTQIYVIDPEP 145

RESULT 4  
US-09-082-593-10  
; Sequence 10, Application US/09082593  
; Patent No. 6180104  
; GENERAL INFORMATION:  
; APPLICANT: DAVIS, MARK M.  
; APPLICANT: HEDRICK, STEPHEN M.  
; TITLE OF INVENTION: T CELL RECEPTOR BETA SUBUNIT  
; FILE REFERENCE: JX1193-195DIV2  
; CURRENT APPLICATION NUMBER: US/09/082,593  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-082-593-10

Query Match 10.4%; Score 106; DB 3; Length 270;  
Best Local Similarity 25.8%; Pred. No. 0.0026;  
Matches 40; Conservative 22; Mismatches 45; Indels 48; Gaps 9;  
QY 3 MLLILIMVHSGCALWVS-----QPPEIRTEGSSAFPCSFNASQGLRAIGSVTW 55  
DB 12 WLL-----NWVNSQQNVQSPESLIVPEGARTSLNCTFSDSASQY----FWW 55  
QY 56 FRDEVVPQKEVR-----NGTPEFRGLAPLASSRFLHDQAELHIRDVRGHDSIYV 108  
DB 56 YRQH--SGKAPKALMSIFSNGEKE-EGRFTIHLNKASLH---FSLHIRDSPDSALYLC 109  
QY 109 RVEVLGLGVGTGNGTRLVVEKEHPQLGAGTVLLIR 143  
DB 110 AVTLVG-----GSGNKLI-----FGTGILLSVK 132

RESULT 5  
US-09-227-595-30  
; Sequence 30, Application US/09227595

```

; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/227,595
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/595,590
; FILING DATE: February 2, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-09-227-595--32

Query Match 10.2%; Score 104.5; DB 4; Length 238;
Best Local Similarity 24.7%; Pred. No. 0.0032;
Matches 43; Conservative 24; Mismatches 70; Indels 37; Gaps 7;

QY 1 MAMWLLILLIM-----VHPGSCALWYSQPEPTLGGSSAPLPCSFNASQRLAIGSVTW 55
Db 1 MGWSCIILFLVATGVH---SAMHVAQPAVLASSRGIAFVCEY--ASPGKATEVRVTV 56
QY 56 FRDEVVPGKEVRNGTPEFRGRGLAPLASSRFLHP-----HQAELHIRDVRGHDAIY 106
Db 57 LRQADSQVTEVCAATYMGNELT-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLY 110
QY 107 VCEVEVL---GLGVGTGNGTRLVV-----EKEHPQLGAGTVLLLRAGFY 147
Db 111 ICKVELMPPYYILGIGNGAQITVAAPSVIFPPSDQLKSGTASVVCLLNIFY 164

RESULT 7
US-08-804-180C-4
; Sequence 4, Application US/08804180C
; Patent No. 6107056
; GENERAL INFORMATION:
; APPLICANT: Martin K. Oaks
; TITLE OF INVENTION: sCTLA-4 and Its Soluble Products
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas M. Wozny
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.50 inch Disk
; COMPUTER: IBM
; OPERATING SYSTEM: DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,180C
; FILING DATE: February 20, 1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas M. Wozny
; REGISTRATION NUMBER: 28,922

```

```

; TELEPHONE: 310/312-9900
; TELEFAX: 310/479-8340
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-067-684-14

Query Match 10.1%; Score 103.5; DB 1; Length 187;
Best Local Similarity 27.7%; Pred.No.0.0029;
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5

QY 18 ALWVSQPEIRTLGSSAFLPCSFNASQGLAIGSVTWPRDEVVPGKEVRNGTPEFRGL 77
Db 1 AMEVAQPAVVLASSRGIA SFCEY-ASPGKATEVRVTVLQADSQVTEVCAATYMGNEL 59
QY 78 APLASSRFLHD-----HQELHTRDVRGHDASIYCRVEVL---GLGVGTGNGTRL 125
Db 60 T-----FLDPSICTGSSGNQVNLTIQGRANDTGLYICKVELMYPPPYILGIGNGTQI 113
QY 126 VVEKEHPQLGAGTVLLRA---GFYAVSFLSVAV 156
Db 114 YVIDPEPCFSDDFILWILAAVSSGLFFYSFLLTAV 148

RESULT 9
US-08-008-898-14
; Sequence 14, Application US/08008898
; Patent No. 5770197
; GENERAL INFORMATION:
; APPLICANT: Linsley, Peter S
; APPLICANT: Ledbetter, Jeffrey A
; APPLICANT: Damle, Nitin K
; APPLICANT: Brady, William
; TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 201 South Lake Avenue, Suite 800
; CITY: Pasadena
; STATE: California
; COUNTRY: United States
; ZIP: 91101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,898
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/723,617
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandel, SaraLynn
; REGISTRATION NUMBER: 31,853
; REFERENCE/DOCKET NUMBER: 7849
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 187 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-008-898-14

```



Query Match 10.1%; Score 103.5; DB 1; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77  
DB 1 AHVQAQPAVVLAASSRGIAFPVCEY-ASPGKATEVRVTVLRLQADSQVTEVCAATYMMGNEL 59

QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125  
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113

QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156  
DB 114 YVIDPEPCPDSDFLLWILAASVSSGLFFYSFLLTAV 148

RESULT 10  
US-08-459-818-14  
; Sequence 14, Application US/08459818  
; Patent No. 5851795  
; GENERAL INFORMATION:  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Damle, Nitin K.  
; APPLICANT: Brady, William  
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant & Gould  
; STREET: 11150 Santa Monica Blvd., Suite 400  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: FastSeq 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/459,818  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adriano, Sarah B.  
; REGISTRATION NUMBER: 34,470  
; REFERENCE/DOCKET NUMBER: 30436.35US02  
; TELEPHONE: 310-445-1140  
; TELEFAX: 310-445-9031  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 187 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-459-818-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77  
DB 1 AHVQAQPAVVLAASSRGIAFPVCEY-ASPGKATEVRVTVLRLQADSQVTEVCAATYMMGNEL 59

QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125  
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113

QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156

DB 114 YVIDPEPCPDSDFLLWILAASVSSGLFFYSFLLTAV 148

RESULT 11  
US-08-889-666-14  
; Sequence 14, Application US/0889666  
; Patent No. 5885579  
; GENERAL INFORMATION:  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Damle, Nitin K.  
; APPLICANT: Brady, William  
; APPLICANT: Kiener, Peter A.  
; TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant & Gould  
; STREET: 11150 Santa Monica Blvd., Suite 400  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/889,666  
; FILING DATE: 08-JUL-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/375390  
; FILING DATE: 18-JAN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Adriano, Sarah B.  
; REGISTRATION NUMBER: 34,470  
; REFERENCE/DOCKET NUMBER: 30436-35US01  
; TELEPHONE: 310-445-1140  
; TELEFAX: 310-445-9031  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 187 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-889-666-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVPVQKEVRNGTPEFRGRL 77  
DB 1 AHVQAQPAVVLAASSRGIAFPVCEY-ASPGKATEVRVTVLRLQADSQVTEVCAATYMMGNEL 59

QY 78 APLASSRFLHD-----HQAEHLHVRDVRGHDASIVYCRVEVL---GLGVGTGNGTRL 125  
DB 60 T-----FLDSDICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPLGIGNGTQI 113

QY 126 VVEKEHPQLGAGTVLLRA-----GFYAVSFLSVAV 156  
DB 114 YVIDPEPCPDSDFLLWILAASVSSGLFFYSFLLTAV 148

RESULT 12  
US-08-465-078-14  
; Sequence 14, Application US/08465078  
; Patent No. 5885796  
; GENERAL INFORMATION:

APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Damle, Nitin K.  
APPLICANT: Brady, William  
APPLICANT: Kiener, Peter A.  
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Blvd., Suite 400  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,078  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375390  
FILING DATE: 18-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-35US01  
TELEPHONE: 310-445-1140  
TELEFAX: 310-445-9031  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-465-078-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSNASQGRLAIGSVTFWFRDEVVPGKEVRNGTPEFRGRL 77  
Db 1 AMHVAQPAVLAASSRGIAFVCEY-ASPGKATEVRVTVLRLQADSQVTEVCAATYMGNEL 59

QY 78 APLASSRLHD-----HQAELHIRDVRGHDASIVYCRVEVL-----GLGVGTGNGTRL 125  
Db 60 T-----FLDPSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMPPPYLGGNGTQI 113

QY 126 VVEKEHPOLGAGTVLLRA-----GFYAVSFLSVAV 156  
Db 114 YVIDPEPCDSDFLLWILAIVSSGLFFYSFLLTAV 148

RESULT 13  
US-08-725-776-14  
Sequence 14, Application US/08725776  
Patent No. 5968510  
GENERAL INFORMATION:  
APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Damle, Nitin K.  
APPLICANT: Brady, William  
APPLICANT: Kiener, Peter A.  
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Blvd., Suite 400

CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/725,776  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375390  
FILING DATE: 18-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-35US01  
TELEPHONE: 310-445-1140  
TELEFAX: 310-445-9031  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-725-776-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;

QY 18 ALWVSQPEIRTELGSSAFPCSNASQGRLAIGSVTFWFRDEVVPGKEVRNGTPEFRGRL 77  
Db 1 AMHVAQPAVLAASSRGIAFVCEY-ASPGKATEVRVTVLRLQADSQVTEVCAATYMGNEL 59

QY 78 APLASSRLHD-----HQAELHIRDVRGHDASIVYCRVEVL-----GLGVGTGNGTRL 125  
Db 60 T-----FLDPSICTGTSSGNQVNLTIQGLRAMDTGLYICKVELMPPPYLGGNGTQI 113

QY 126 VVEKEHPOLGAGTVLLRA-----GFYAVSFLSVAV 156  
Db 114 YVIDPEPCDSDFLLWILAIVSSGLFFYSFLLTAV 148

RESULT 14  
US-08-488-062-14  
Sequence 14, Application US/08488062  
Patent No. 5977318  
GENERAL INFORMATION:  
APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Damle, Nitin K.  
APPLICANT: Brady, William  
APPLICANT: Kiener, Peter A.  
TITLE OF INVENTION: CTLA4 Receptor and Uses Thereof  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Blvd., Suite 400  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,062  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/375390  
FILING DATE: 18-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-35US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310-445-9031  
TELEFAX: 310-445-9031  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-062-14

Query Match 10.1%; Score 103.5; DB 2; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;  
QY 18 ALWVSQPPEIRTLGSSAFPCSFNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGL 77  
Db 1 AMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVTVLROADSQVTEVCAATYMMGNEL 59  
QY 78 APLASSRFLHD-----HQAELHIRDVRGHDASIYVCRVEVL---GLGVGTGNGTRL 125  
Db 60 T-----FLDDSICTGTSSGNQVNTIQGLRAMDTGLYICKVELMYPYPYVLGIGNGTQI 113  
QY 126 VVEKEHPQAGTVLLRA-----GFYAVSFLSVAV 156  
Db 114 YVIDPEPCDSDFLWILAAVSSGLFFYSFLLTAV 148

## RESULT 15

US-08-228-208A-14  
Sequence 14, Application US/08228208A  
Patent No. 6090914  
GENERAL INFORMATION:  
APPLICANT: Linsley, Peter S.  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Damle, Nitin K.  
APPLICANT: Brady, William  
APPLICANT: Wallace, Philip M.  
TITLE OF INVENTION: CTLA4/CD28lg HYBRID FUSION  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant & Gould  
STREET: 11150 Santa Monica Boulevard, Suite 400  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/228,208A  
FILING DATE: 15-APR-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/008,898  
FILING DATE: 22-JAN-1993  
APPLICATION NUMBER: 07/723,617  
FILING DATE: 27-JUN-1991  
ATTORNEY/AGENT INFORMATION:

NAME: Adriano, Sarah B.  
REGISTRATION NUMBER: 34,470  
REFERENCE/DOCKET NUMBER: 30436-30US01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310 445-1140  
TELEFAX: 310 445-9031  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 187 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-228-208A-14

Query Match 10.1%; Score 103.5; DB 3; Length 187;  
Best Local Similarity 27.7%; Pred. No. 0.0029;  
Matches 43; Conservative 20; Mismatches 69; Indels 23; Gaps 5;  
QY 18 ALWVSQPPEIRTLGSSAFPCSFNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGL 77  
Db 1 AMHVAQPAVVLASSRGIAFVCEY-ASPGKATEVTVLROADSQVTEVCAATYMMGNEL 59  
QY 78 APLASSRFLHD-----HQAELHIRDVRGHDASIYVCRVEVL---GLGVGTGNGTRL 125  
Db 60 T-----FLDDSICTGTSSGNQVNTIQGLRAMDTGLYICKVELMYPYPYVLGIGNGTQI 113  
QY 126 VVEKEHPQAGTVLLRA-----GFYAVSFLSVAV 156  
Db 114 YVIDPEPCDSDFLWILAAVSSGLFFYSFLLTAV 148

Search completed: February 26, 2004, 12:16:14  
Job time : 37.7905 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 51.2467 Seconds  
(without alignments)  
738.822 Million cell updates/sec

Title: US-10-036-444-4  
Perfect score: 632  
Sequence: 1 LNWSPPEIRTEGSSAFLP.....TGNTRLVVEKHPQLGAGT 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25:  
1: sp archaea:  
2: sp bacteria:  
3: sp fungi:  
4: sp human:  
5: sp invertebrate:  
6: sp mammal:  
7: sp mhc:  
8: sp organelle:  
9: sp phase:  
10: sp plant:  
11: sp rodent:  
12: sp virus:  
13: sp vertebrate:  
14: sp unclassified:  
15: sp xvirus:  
16: sp bacteriap:  
17: sp archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	632	100.0	177	014930	014930 homo sapien
2	632	100.0	190	014932	014932 homo sapien
3	632	100.0	201	014931	014931 homo sapien
4	599	94.8	176	095JB8	Q95JB8 macaca fasc
5	599	94.8	180	Q8MJ02	Q8MJ02 macaca mula
6	476.5	75.4	152	095668	Q95668 homo sapien
7	476.5	75.4	165	095669	Q95669 homo sapien
8	476.5	75.4	176	095667	Q95667 homo sapien
9	466	73.7	115	Q8MJ00	Q8MJ00 macaca mula
10	454.5	71.9	151	Q8MJ01	Q8MJ01 macaca mula
11	428	67.7	192	11 Q8CFD9	Q8CFD9 rattus norv
12	424	67.1	192	11 Q8QWR8	Q8QWR8 rattus norv
13	420	66.5	192	11 Q8CG11	Q8CG11 rattus norv
14	108.5	17.2	449	5 Q9NKA5	Q9NKA5 drosophila
15	105.5	16.7	235	11 Q99M11	Q99M11 mus musculu
16	103	16.3	329	4 Q8N225	Q8N225 homo sapien

17	103	16.3	1340	4 Q8NDA2	Q8NDA2 homo sapien
18	101.5	16.1	221	6 Q28090	Q28090 bos taurus
19	100.5	15.9	221	6 Q97631	Q97631 ovis aries
20	96.5	15.3	350	5 Q9VFU7	Q9VFU7 drosophila
21	94	14.9	526	4 Q9H458	Q9H458 homo sapien
22	93.5	14.8	274	11 Q80YU5	Q80YU5 mus musculu
23	92.5	14.6	223	11 Q7TWX1	Q7TWX1 mus musculu
24	92	14.6	174	11 Q921A7	Q921A7 rattus norv
25	91.5	14.5	223	6 Q9TT02	Q9TT02 canis famil
26	91.5	14.5	223	6 Q9GKP2	Q9GKP2 canis famil
27	91.5	14.5	223	11 Q62859	Q62859 rattus norv
28	90.5	14.3	272	11 Q70356	Q70356 mus musculu
29	90	14.2	524	11 Q921K7	Q921K7 mus musculu
30	90	14.2	700	11 Q7TSU7	Q7TSU7 mus musculu
31	89.5	14.2	137	4 Q95653	Q95653 homo sapien
32	89	14.1	108	13 Q8JU25	Q8JU25 orectolobus
33	88.5	14.0	223	6 Q9XTA1	Q9XTA1 felis silve
34	88.5	14.0	223	6 Q9XSY7	Q9XSY7 felis silve
35	88.5	14.0	244	6 Q7YRD9	Q7YRD9 macaca fasc
36	88.5	14.0	247	6 Q9BGS7	Q9BGS7 macaca fasc
37	88	13.9	178	11 Q9D1U4	Q9D1U4 mus musculu
38	88	13.9	179	11 Q921X1	Q921X1 mus musculu
39	88	13.9	523	11 Q8K2H7	Q8K2H7 mus musculu
40	88	13.9	946	13 Q07153	Q07153 torpedo cal
41	87.5	13.8	141	4 Q9NU66	Q9NU66 homo sapien
42	87.5	13.8	160	4 Q8TDA6	Q8TDA6 homo sapien
43	87.5	13.8	176	4 Q9NU65	Q9NU65 homo sapien
44	87.5	13.8	183	4 Q9NU64	Q9NU64 homo sapien
45	87.5	13.8	224	4 Q96KV1	Q96KV1 homo sapien

## ALIGNMENTS

## RESULT 1

014930 PRELIMINARY; PRT; 177 AA.  
ID 014930  
AC 014930; 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE 1C7 precursor (1C7 protein).  
GN 1C7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=96422187; PubMed=8824804;  
RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weisman S.M.;  
RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";  
RL Genomics 31:215-222(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX Nalabolu S.R., Raghunathan A., Weisman S.M.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93272029; PubMed=8499947;  
RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;  
RT "Dense Alu clustering and a potential new member of the NF kappa B  
family within a 90 kilobase HLA class III segment.";  
RL Nat. Genet. 3:137-145(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96215741; PubMed=8629302;  
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
RA Wallace A.F., Russell M.E.;  
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage

RT molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96005657; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwiersina H.,  
 RA Weiss E.H.;  
 RN "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208891; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.,  
 RA "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95316093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RA "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RA "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RA "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53(1991).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RA "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
 RA Van Zeveren A., Coppieers W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RA "The BAT1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family.";  
 RL Genomics 26:210-218(1995).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RA "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing.";

RL Genomics 45:591-600(1997).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149885; PubMed=9480751;  
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IKB1 and MICA genes at the centromeric end of the HLA class I  
 RT region.";  
 RL Genomics 47:372-382(1998).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RA Shiina S., Tamiya G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF031136; AAB86578.1; --  
 DR EMBL; Y14768; CAA75064.1; --  
 DR EMBL; AP000505; BAB83393.1; --  
 DR Genew; HGNC:4189; LST1.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0003793; P:defense/immunity protein activity; TAS.  
 DR GO; GO:0006968; P:cellular defense response; TAS.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; ig\_1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG Like; 1.  
 KW Signal.  
 FT SIGNAL 1 61 POTENTIAL.  
 FT CHAIN 62 177 1C7.  
 SQ SEQUENCE 177 AA; 19237 MW; DD5EC96FOAB2DC66 CRC64;  
 Query Match 100.0%; Score 632; DB 4; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-60;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LWSQPPEIRTLGSSAFPLPCSFNASQGLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
 DB 19 LWSQPPEIRTLGSSAFPLPCSFNASQGLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78  
 QY 61 PLASRFLHDHQAELHIRDVRGHDASIIYVCRVEVLGLGVGTGNGTRLVWEKEHPQLGAGT 120  
 DB 79 PLASRFLHDHQAELHIRDVRGHDASIIYVCRVEVLGLGVGTGNGTRLVWEKEHPQLGAGT 138  
 RESULT 2  
 C14932  
 ID C14932 PRELIMINARY; PRT; 190 AA.  
 AC C14932;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Lc7 precursor (Natural killer cell receptor).  
 GN Lc7 OR NKP30.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary;  
 RX MEDLINE=96422187; PubMed=8824804;  
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;  
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";  
 RL Genomics 31:215-222(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.



RESULT 3

O14931 PRELIMINARY; PRT; 201 AA.

AC O14931;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)

DE 01-OCT-2003 (TrEMBLrel. 25, last annotation update)

DE 1C7 precursor (NCR3 protein).

GN 1C7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Spleen;

FX MEDLINE=96422187; PubMed=8824804;

RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;

RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";

RL Genomics 31:215-222(1996).

[2]

RN SEQUENCE FROM N.A.

RP TISSUE=Spleen;

RA Nalabolu S.R., Raghunathan A., Weissman S.M.;

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

[3]

RN SEQUENCE FROM N.A.

RP Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,

RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,

RA Lasky S., Hood L.;

RT "Sequence of the human major histocompatibility complex class III

RT region.";

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.

[4]

RN SEQUENCE FROM N.A.

RP MEDLINE=9272029; PubMed=8499947;

RA Iris P., Bougueret L., Prieur S., Caterina D., Primas G., Perrot V.,

RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.,

RT "Dense Alu clustering and a potential new member of the NP kappa B

RT family within a 90 kilobase HLA class III segment.";

RL Nat. Genet. 3:137-145(1993).

[5]

RN SEQUENCE FROM N.A.

RP MEDLINE=96215741; PubMed=8629302;

RA Urans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,

RA Wallace A.F., Russell M.E.;

RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage

RT molecule expressed in transplanted human hearts.";

RL Transplantation 61:1387-1392(1996).

[6]

RN SEQUENCE FROM N.A.

RP MEDLINE=96006565; PubMed=7590964;

RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,

RA Weiss E.H.;

RT "Cloning and genomic characterization of LST1: a new gene in the human

RT TNF region.";

RL Immunogenetics 42:315-322(1995).

[7]

RN SEQUENCE FROM N.A.

RP MEDLINE=93208881; PubMed=7916655;

RA Browning J.L., Ngan-ek A., Lawton P., DeMarinis J., Tizard R.,

RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.E.;

RT "Lymphotoxin beta, a novel member of the TNF family that forms a

RT heteromeric complex with lymphotoxin on the cell surface.";

RL Cell 72:847-856(1993).

[8]

RN SEQUENCE FROM N.A.

RP MEDLINE=86016093; PubMed=2995927;

RA Pennin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,

RA Pennin D., Goeddel D.V., Gray P.W.;

RT "Human lymphotoxin and tumor necrosis factor genes: structure,

RT homology and chromosomal localization.";

RL Nucleic Acids Res. 13:6361-6373(1985).

[9]

RN SEQUENCE FROM N.A.

RP MEDLINE=91086846; PubMed=1670638;

RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,

RA Riethmuller G., Weiss E.H.;

RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an

RT NcoI polymorphism in the first intron of the human TNF-beta gene

RT correlates with a variant amino acid in position 26 and a reduced

RT level of TNF-beta production.";

RL J. Exp. Med. 173:209-219(1991).

[10]

RN SEQUENCE FROM N.A.

RP MEDLINE=91139175; PubMed=1671667;

RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;

RT "Haplotypic polymorphisms of the TNF gene.";

RL Immunogenetics 33:50-53(1991).

[11]

RN SEQUENCE FROM N.A.

RP MEDLINE=94362679; PubMed=8081366;

RA Albertella M.R., Campbell D.R.;

RT "Characterization of a novel gene in the human major

RT histocompatibility complex that encodes a potential new member of the

RT I kappa B family of proteins.";

RL Hum. Mol. Genet. 3:793-799(1994).

[12]

RN SEQUENCE FROM N.A.

RP MEDLINE=95324911; PubMed=7601445;

RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,

RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,

RA Strominger J., Spies T.;

RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative

RT nuclear RNA helicase of the DEAD family.";

RL Genomics 26:210-218(1995).

[13]

RN SEQUENCE FROM N.A.

RP MEDLINE=20132445; PubMed=10668961;

RA Neville M.J., Campbell R.D.;

RT "Alternative splicing of the LST-1 gene located in the major

RT histocompatibility complex on human chromosome 6.";

RL DNA Seq. 8:155-160(1997).

[14]

RN SEQUENCE FROM N.A.

RP MEDLINE=98035883; PubMed=9367684;

RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidle U.,

RA Weiss E.H.;

RT "Complex expression pattern of the TNF region gene LST1 through

RT differential regulation, initiation, and alternative splicing.";

RL Genomics 45:591-600(1997).

[15]

RN SEQUENCE FROM N.A.

RP MEDLINE=98149985; PubMed=9480751;

RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,

RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,

RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,

RA Kimura M., Inoko H.;

RT "Nucleotide sequencing analysis of the 146-kilobase segment around the

RT Ikb1 and MICA genes at the centromeric end of the HLA class I

RT region.";

RL Genomics 47:372-382(1998).

[16]

RN SEQUENCE FROM N.A.

RP TISSUE=Blood;

FX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.I., Scheetz T.E.,

RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Maltby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,



RA Villalon D.K., Muzny D.M., Sodergren E.J., Liu X., Gibbs R.A.,  
RA Pahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences",  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [17]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Blood;  
RA Strausberg R.;  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR ENBL; AF031137; AAB86579.1; -  
DR ENBL; AF129756; AAD18088.1; -  
DR ENBL; Y14768; CAAV5063.1; -  
DR ENBL; BC052582; AAB52582.1; -  
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0006954; P:inflammatory response; NAS.  
DR InterPro; IPR003599; Ig-like.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
KW SIGNAL.  
FT SIGNAL.  
FT CHAIN 1 61 POTENTIAL.  
FT CHAIN 62 201 1C7.  
SQ SEQUENCE 201 AA; 21593 MW; 2855AE4D6902D429 CRC64;  
Query Match 100.0%; Score 632; DB 4; Length 201;  
Best Local Similarity 100.0%; Pred. No. 2.6e-60;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
DB 19 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78  
QY 61 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 120  
DB 79 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 138  
RESULT 4  
Q95JB8 PRELIMINARY; PRT; 176 AA.  
AC Q95JB8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Nkp30 (Nkp30v1).  
GN NCR.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Macaca.  
OX NCBI\_TaxID=9541; 9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.fascicularis; TISSUE=Lymphoid;  
RA Rizzi M., Blasson R.;  
RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in  
RT Macaca fascicularis lymphoid cells",  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.mulatta;  
RA LaBonte M.L., Miller J., Letvin N.L.;  
RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification  
RT of Nkp46SD and Nkp30S.",  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR ENBL; AY035214; AAK63116.1; -  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;  
Query Match 94.8%; Score 599; DB 6; Length 180;  
Best Local Similarity 95.0%; Pred. No. 8.5e-57;  
Matches 114; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
DB 19 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78  
QY 61 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 120  
DB 79 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 138  
RESULT 6  
Q95668 PRELIMINARY; PRT; 152 AA.  
AC Q95668;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE 1C7d.  
GN 1C7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RT of Nkp46SD and Nkp30S.",

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR ENBL; AJ278399; CAC41081.1; -  
DR ENBL; AY035215; AAK63117.1; -  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;  
Query Match 94.8%; Score 599; DB 6; Length 176;  
Best Local Similarity 95.0%; Pred. No. 8.3e-57;  
Matches 114; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
DB 19 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78  
QY 61 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 120  
DB 79 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 138  
RESULT 5  
Q8MJ02 PRELIMINARY; PRT; 180 AA.  
AC Q8MJ02;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Nkp30.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA LaBonte M.L., Miller J., Letvin N.L.;  
RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification  
RT of Nkp46SD and Nkp30S.",  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR ENBL; AY035214; AAK63116.1; -  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;  
Query Match 94.8%; Score 599; DB 6; Length 180;  
Best Local Similarity 95.0%; Pred. No. 8.5e-57;  
Matches 114; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
DB 19 LWSQPEIIRTLGSSAFPCSNASQRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78  
QY 61 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 120  
DB 79 PLASSRFLHDQAEHLHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 138  
RESULT 6  
Q95668 PRELIMINARY; PRT; 152 AA.  
AC Q95668;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE 1C7d.  
GN 1C7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris P., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NFKappaB  
 RT family within a 90 kilobase HLA class III segment."; Nat.  
 RL Nat. Genet. 3:137-145(1993).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8623902;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.B., Arcaci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RA "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts."; Transplantation  
 RL 61:1387-1392(1996).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=96008565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 RA Weiss E.H.;  
 RA "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region."; Immunogenetics 42:315-322(1995).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=92208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RA "Lymphotoxin-beta: A new member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface."; Cell  
 RL 72:847-856(1993).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RA "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization."; Nucleic Acids Res. 13:6361-6373(1985).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RA "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
 RT Level of TNF-beta Production."; J. Exp. Med. 173:209-219(1991).  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=1671667; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RA "Haplotypic polymorphisms of the TNFB gene."; Immunogenetics 33:50-53(1991).  
 RN [8]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RA "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins."; Hum. Mol. Genet. 3:793-799(1994).  
 RN [9]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vaiman M.,  
 RA van Zeveren A., Coppieers W., Van de Weghe A., Souquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RA "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
 RT Nuclear RNA Helicase of the D-E-A-D Family.";

Genomics 26:210-218(1995).  
 RN [10]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RA "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6."; DNA Seq. 8:155-160(1997).  
 RN [11]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
 RA Weiss E.H.;  
 RA "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing."; Genomics 45:591-600(1997).  
 RN [12]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RA "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IkBL and MICA genes at the centromeric end of the HLA class I  
 RT region."; Genomics 47:372-382(1998).  
 RN [13]  
 RN EMEL; Y14768; CAA75067.1; -;  
 DR GO; GO:0003793; P:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG-like.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS00835; IG Like; 1.  
 SQ SEQUENCE 152 AA; I6393 MW; 42718746451F9ADC CRC64;  
 Query Match 75.4%; Score 476.5; DB 4; Length 152;  
 Best Local Similarity 79.2%; Pred. No. 1.3e-43;  
 Matches 95; Conservative 0; Mismatches 0; Indels 25; Gaps 1;  
 QY 1 LWSQPPPIRTLEGSSAFPLCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
 Db 19 LWSQPPPIRTLEGSSAFPLCSFNASQGRLAIGSVTWFRDEVVPGKE----- 65  
 QY 61 PLASRFUHDQAEHLHVDVRGHDASIYVCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 120  
 Db 66 -----AELHVDVRGHDASIYVCRVEVLGLGVGTGNGTRLVWKEHPQLGAGT 113  
 RESULT 7  
 O95669  
 ID O95669 PRELIMINARY; PRT; 165 AA.  
 AC O95669;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 1C7F.  
 GN 1C7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris P., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NFKappaB  
 RT family within a 90 kilobase HLA class III segment."; Nat.  
 RL Nat. Genet. 3:137-145(1993).  
 RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=96215741; PubMed=8629302;  
RA Utans U., Quiet W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
RA Wallace A.F., Russell M.E.;  
RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
RT molecule expressed in transplanted human hearts.";  
RL Transplantation 61:1387-1392(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96006565; PubMed=7590964;  
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
RA Weiss E.H.;  
RT "Cloning and genomic characterization of LST1: a new gene in the human  
RT TNF region";  
RL Immunogenetics 42:315-322(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9320881; PubMed=7916655;  
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
RT "Lymphotoxin-beta: A new member of the TNF family that forms a  
RT heteromeric complex with lymphotoxin on the cell surface.";  
RL Cell 72:847-856(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86016093; PubMed=2995927;  
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
RA Pennica D., Goeddel D.V., Gray P.W.;  
RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
RT homology and chromosomal localization.";  
RL Nucleic Acids Res. 13:6361-6373(1985).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91086846; PubMed=1670638;  
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
RA Rietmueller G., Weiss E.H.;  
RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
RT Level of TNF-beta Production.";  
RL J. Exp. Med. 173:209-219(1991).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91139175; PubMed=1671667;  
RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
RT "Haplotypic polymorphisms of the TNFS gene.";  
RL Immunogenetics 33:50-53(1991).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94362679; PubMed=8081366;  
RA Albertella M.R., Campbell D.R.;  
RT "Characterization of a novel gene in the human major  
RT histocompatibility complex that encodes a potential new member of the  
RT I kappa B family of proteins.";  
RL Hum. Mol. Genet. 3:793-799(1994).  
RN [9]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95324911; PubMed=7601445;  
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
RA Strominger J., Spies T.;  
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
RT Nuclear RNA Helicase of the D-E-A-D Family.";  
RL Genomics 26:210-218(1995).  
RN [10]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20132445; PubMed=10668961;  
RA Neville M.J., Campbell R.D.;  
RT "Alternative splicing of the LST-1 gene located in the major  
RT histocompatibility complex on human chromosome 6.";  
RL DNA Seq. 8:155-160(1997).  
RN [11]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96006565; PubMed=7590964;  
RX MEDLINE=98035883; PubMed=9367684;  
de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
Weiss E.H.;  
RT "Complex expression pattern of the TNF region gene LST1 through  
RT differential regulation, initiation, and alternative splicing.";  
RL Genomics 45:591-600(1997).  
RN [12]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98149985; PubMed=9480751;  
RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura I.,  
RA Kimura M., Inoko H.;  
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
RT IKBL and MICA genes at the centromeric end of the HLA class I  
RT region.";  
RL Genomics 47:372-382(1998).  
DR EMBL; Y14768; CAA75068.1; ...  
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0006954; P:inflammatory response; NAS.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;  
Query Match 75.4%; Score 476.5; DB 4; Length 165;  
Best Local Similarity 79.2%; Pred. No. 1.4e-43;  
Matches 95; Conservative 0; Mismatches 0; Indels 25; Gaps 1;  
QY 1 LWSQPPPIRTLEGSSAFPLPCSFNASQGLAIGSVTWFRDEVVPGKEVRNGTTPFRGRLA 60  
Db 19 LWSQPPPIRTLEGSSAFPLPCSFNASQGLAIGSVTWFRDEVVPGKE----- 65  
QY 61 PLASSRFLHQAELHINDVRGHDASIVYCRVEVLGLVGVTGNGTRLVVKEKHPQLGAGT 120  
Db 66 -----AELHINDVRGHDASIVYCRVEVLGLVGVTGNGTRLVVKEKHPQLGAGT 113  
RESULT 8  
O95667 PRELIMINARY; PRT; 176 AA.  
AC O95667  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE 1C7e.  
GN 1C7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93272029; PubMed=8499947;  
RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;  
RT "Dense Alu clustering and a potential new member of the Nfkapab  
RT family within a 90 kilobase HLA class III segment.";  
RL Nat. Genet. 3:137-145(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96215741; PubMed=8629302;  
RA Utans U., Quiet W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
RA Wallace A.F., Russell M.E.;  
RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
RT molecule expressed in transplanted human hearts.";  
RL Transplantation 61:1387-1392(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96006565; PubMed=7590964;

RA Holzinger I., de Baey A., Messer G., Kick G., Zwierniza H.,  
RA Weiss E.H.;  
RA "Cloning and genomic characterization of LST1: a new gene in the human  
RA TNF region.",  
RL Immunogenetics 42:315-322(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93208881; PubMed=7916655;  
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.,  
RA "Lymphotoxin-beta: A new member of the TNF family that forms a  
RA heteromeric complex with lymphotoxin on the cell surface.",  
RL Cell 72:847-856(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86016093; PubMed=2995927;  
RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
RA Pennica D., Goeddel D.V., Gray P.W.,  
RA "Human lymphotoxin and tumor necrosis factor genes: structure,  
RA homology and chromosomal localization.",  
RL Nucleic Acids Res. 13:6361-6373(1985).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=1086846; PubMed=1670638;  
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
RA Riethmuller G., Weiss E.H.,  
RA "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
RA Neol Polymorphism in the First Intron of the Human TNF-beta Gene  
RA Correlates with a Variant Amino Acid in Position 26 and a Reduced  
RA Level of TNF-beta Production.",  
RL J. Exp. Med. 173:209-219(1991).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91139175; PubMed=1671667;  
RA Abraham I.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.,  
RA "Haplotypic polymorphisms of the TNFB gene.",  
RL Immunogenetics 33:50-53(1991).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94362679; PubMed=8081366;  
RA Albertella M.R., Campbell D.R.,  
RA "Characterization of a novel gene in the human major  
RA histocompatibility complex that encodes a potential new member of the  
RA I kappa B family of proteins.",  
RL Hum. Mol. Genet. 3:793-799(1994).  
RN [9]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95324911; PubMed=7601445;  
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
RA Van Zeveren A., Coppiepers W., Van de Weghe A., Bouquet Y., Choy W.,  
RA Strominger J., Spies T.,  
RA "The B2M Gene in the MHC Encodes an Evolutionarily Conserved Putative  
RA Nuclear RNA Helicase of the D-E-A-D Family.",  
RL Genomics 26:210-218(1995).  
RN [10]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20132445; PubMed=10668961;  
RA Neville M.J., Campbell R.D.,  
RA "Alternative splicing of the LST-1 gene located in the major  
RA histocompatibility complex on human chromosome 6.",  
RL DNA Seq. 8:155-160(1997).  
RN [11]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98035883; PubMed=9367684;  
RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidle U.,  
RA Weiss E.H.,  
RA "Complex expression pattern of the TNF region gene LST1 through  
RA differential regulation, initiation, and alternative splicing.",  
RL Genomics 45:591-600(1997).  
RN [12]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98149985; PubMed=9480751;  
RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,

RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
RA Kimura M., Inoko H.,  
RA "Nucleotide sequencing analysis of the 146-kilobase segment around the  
RA IKB1 and MICA genes at the centromeric end of the HLA class I  
RA region.",  
RL Genomics 47:372-382(1998).  
RX EMBL; Y14768; CAA75066.1; -.  
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0006954; P:inflammatory response; NAS.  
DR InterPro; IPR003599; IG-like.  
DR InterPro; IPR007110; IG-like.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
SQ SEQUENCE 176 AA; 18749 MW; 162BB775DA2BCD35 CRC64;  
Query Match 75.4%; Score 476.5; DB 4; Length 176;  
Best Local Similarity 79.2%; Pred. No. 1.5e-43;  
Matches 95; Conservative 0; Mismatches 0; Indels 25; Gaps 1;  
QY 1 LWSQPPPEIRTLGSSAFLPCSFNASQGLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
DB 19 LWSQPPPEIRTLGSSAFLPCSFNASQGLAIGSVTWFRDEVVPGKE----- 65  
QY 61 PLASRFLRDHQAEHLHVRGHDASIVYCRVEVLGLVGTGTGTRLVWEKEHPQLGAGT 120  
DB 66 -----AELHVRGHDASIVYCRVEVLGLVGTGTGTRLVWEKEHPQLGAGT 113  
RESULT 9  
Q8MJ00 PRELIMINARY; PRT; 115 AA.  
ID Q8MJ00  
AC Q8MJ00;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE NKP30S.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA LaBonte M.L., Miller J., Letvin N.L.,  
RA "Molecular cloning of rhesus monkey Nkp30 and identification  
RA of Nkp30 and Nkp30S.",  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035217; AAK63119.1; -.  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
SQ SEQUENCE 115 AA; 12871 MW; 1D34CC0B986DE9F CRC64;  
Query Match 73.7%; Score 466; DB 6; Length 115;  
Best Local Similarity 94.6%; Pred. No. 1.2e-42;  
Matches 88; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 LWSQPPPEIRTLGSSAFLPCSFNASQGLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
DB 19 LWSQPPPEIRTLGSSAFLPCSFNASQGLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78  
QY 61 PLASRFLRDHQAEHLHVRGHDASIVYCRVE 93  
DB 79 PLSSRFLRDHQAEHLHVRGHDAGIYVCRVE 111  
RESULT 10  
Q8MJ01

```

ID QBMJ01 PRELIMINARY; PRT; 151 AA.
AC QBMJ01; 2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE NKP30v1ED.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RT LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification
RT of Nkp46SD and Nkp30S."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035216; AAK63118.1; -.
DR InterPro; IPR003599; IG.
DR Pfam; PF00047; IG; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;

Query Match 71.9%; Score 454.5; DB 6; Length 151;
Best Local Similarity 75.8%; Pred. No. 3.1e-41;
Matches 91; Conservative 1; Mismatches 3; Indels 25; Gaps 1;

QY 1 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
DB 66 -----AELHIWDVRGHDAGIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 113

RESULT 11
Q8CFD9 PRELIMINARY; PRT; 192 AA.
AC Q8CFD9;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE lc7 protein precursor.
GN lc7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RT Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor lc7."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430418; CAD23066.1; -.
DR InterPro; IPR003599; IG.
DR Pfam; PF00047; IG; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Signal.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 192 AA; 20500 MW; 7FDD5AB252D239C7 CRC64;

Query Match 67.7%; Score 428; DB 11; Length 192;
Best Local Similarity 65.8%; Pred. No. 3e-38;
Matches 79; Conservative 19; Mismatches 22; Indels 0; Gaps 0;

```

```

QY 1 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120
DB 79 SFSASQFIRGHKAGLLIQDISHDARIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 138

RESULT 12
Q8QWM8 PRELIMINARY; PRT; 192 AA.
ID Q8QWM8;
AC Q8QWM8;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE NKP30.
GN NKP30.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=LEW;
RC STRAIN=LEW; PubMed=12180816;
RX MEDLINE=22168131; Obara H., Ogura Y., Martinez O.M., Krams S.M.;
RA Hsieh C.L., Obara H., Ogura Y., Martinez O.M., Krams S.M.;
RT "NK cells and transplacental."
RL Transpl. Immunol. 9:111-114(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW;
RA Hsieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RA Martinez O.M., Krams S.M.;
RT "Identification, Cloning, and Characterization of a Novel Rat NK
RT Receptor, rNKp30: a Molecule Expressed in Liver Allografts."
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY273824; AAP13457.1; -.
DR InterPro; IPR003599; IG.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 192 AA; 20498 MW; 7FDF58B245C52377 CRC64;

Query Match 67.1%; Score 424; DB 11; Length 192;
Best Local Similarity 67.8%; Pred. No. 8.2e-38;
Matches 78; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

QY 1 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60
DB 19 LWSQPPPEIRTELGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78
QY 61 PLASSRFLHDHQAELHIRDVRGHDASIVYCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 115
DB 79 SFSASQFIRGHKAGLLIQDISHDARIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 133

RESULT 13
Q8CG11 PRELIMINARY; PRT; 192 AA.
ID Q8CG11;
AC Q8CG11;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE NK receptor lc7 precursor.
GN lc7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RT STRAIN=BN;
RC STRAIN=BN;

```

RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,  
RA Butcher G.W.;  
RT "Molecular characterization of the novel rat NK receptor lc7.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ430419; CAD23067.2;  
DR EMBL; AJ430420; CAD23067.2; JOINED.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; Ig; 1.  
DR SWART; SM00409; Ig; 1.  
DR PROSITE; PS00835; IG\_LIKE; 1.  
KW S:signal; Receptor.  
FT SIGNAL 1 19 POTENTIAL.  
SQ SEQUENCE 192 AA; 20470 MW; 439AD7A3APBE6DC0 CRC64;  
  
Query Match 66.5%; Score 420; DB 11; Length 192;  
Best Local Similarity 67.8%; Pred. No. 2.2e-37;  
Matches 78; Conservative 17; Mismatches 20; Indels 0; Gaps 0;  
  
QY 1 LWSQPPPEIRTELEGSAFLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
Db 19 VVVSQPPPEIRAQGTGTASLPCSFNASRGAIGSATWYQDKVAPGMELSNVTIPGFRGVA 78  
  
QY 61 PLASSRFLHQAEHLRDVRGHDASIYVCRVEVLGVGTGNGTGLVVEKEHPQ 115  
Db 79 SFSVSQPIRHKAGLLIQQSHDARIYVCRVEVLGVGTGNGTGLVVEKEPPQ 133  
  
RESULT 14  
Q9NKA5 PRELIMINARY; PRT; 449 AA.  
ID Q9NKA5  
AC Q9NKA5  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN BG:DS01219.3 OR CG15275.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RX MEDLINE=99403001; PubMed=10471707;  
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,  
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,  
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,  
RA Palazzolo M., Reese M.G., Spradling A., Teang G., Wan K., Whitelaw K.,  
RA Celniker S., Rubin G.M.;  
RT "An exploration of the sequence of a 2.9-Mb region of the genome of  
RT Drosophila melanogaster: the Adh region.";  
RL Genetics 153:179-219 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,  
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,  
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,  
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,  
RA Lewis S., Li P., Lonotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,  
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,  
RA Sethi H., Snir E., Svirskaas R.R., Wan K.H., Weinburg T., Zhang R.,  
RA Zieran L.L., Rubin G.M.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003411; AAP44902.1;  
DR FlyBase; FBgn028914; BG:DS01219.3.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00409; Ig; 1.

DR PROSITE; PS00835; IG\_LIKE; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 449 AA; 50945 MW; 162A631C0C172154 CRC64;  
  
Query Match 17.2%; Score 108.5; DB 5; Length 449;  
Best Local Similarity 33.0%; Pred. No. 0.0034;  
Matches 33; Conservative 11; Mismatches 35; Indels 21; Gaps 5;  
  
QY 8 EIRTELEGSAFLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 54  
Db 32 DIDAVEGKSVSLPCPTEPLDNTVM--VLWFRNAGIPLYSFVDRKESREQPRHWSAPZ 89  
  
QY 55 FRGLAPLASSRFLHQAEHLRDVRGHDASIYVCRVE 93  
Db 90 VFG-----SRAKTFDSQPATLEIKDKRHDOGIYRCRD 124  
  
RESULT 15  
Q99M11 PRELIMINARY; PRT; 235 AA.  
ID Q99M11  
AC Q99M11  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognath.; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC02129; AA02129.1; -.  
DR HSSP; P01703; 7FAB.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; IG\_MHC.  
DR InterPro; IPR003596; IG\_v.  
DR Pfam; PF00047; Ig; 2.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 2.  
DR PROSITE; PS00290; IG\_MHC; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 235 AA; 25403 MW; 39807BF65782A3FB CRC64;  
  
Query Match 16.7%; Score 105.5; DB 11; Length 235;  
Best Local Similarity 28.8%; Pred. No. 0.0033;  
Matches 34; Conservative 22; Mismatches 43; Indels 19; Gaps 5;  
  
QY 1 LWSQPPPEIRTELEGSAFLPCSFNASQGRLAIGSVTWFRD-----EVVPGKEVRNG-- 51  
Db 21 LVLTQPSVSTSLGSTAKLPC--KASTGNIGDSYVWYQQYMGKSPNTNIYGDRLRPSGV 78  
  
QY 52 TPEFRGLAPLASSRFLHQAEHLRDVRGHDASIYVCRVEVLGVGTGNGTGLV 109  
Db 79 SDRFSGSIDSSNSAF-----LTIONVQADDEADYICQSYSGIRV-FGGTKLTV 128  
  
Search completed: February 26, 2004, 12:13:42  
Job time : 52.2467 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 74.4828 seconds  
(without alignments)  
455.215 Million cell updates/sec

Title: US-10-036-444-4  
Perfect score: 632  
Sequence: 1 LWSQPPETLTGSSAFLP.....TNGTGLVVEKHPQLGACT 120

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseq1980s:\*  
2: Geneseq1990s:\*  
3: Geneseq2000s:\*  
4: Geneseq2001s:\*  
5: Geneseq2002s:\*  
6: Geneseq2003as:\*  
7: Geneseq2003bs:\*  
8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	100.0	120	4 AAE02771	Aae02771 Human Nkp
2	632	100.0	177	2 AAY06402	Aay06402 Human B-c
3	632	100.0	190	2 AAY06401	Aay06401 Human B-c
4	632	100.0	190	4 AAE02769	Aae02769 Human Nkp
5	632	100.0	201	2 AAY06403	Aay06403 Human B-c
6	617	97.6	135	5 AAE19109	Aae19109 Human Nkp
7	617	97.6	369	5 AAE19110	Aae19110 Human Nkp
8	118.5	18.8	139	5 AAU75565	Aau75565 Murine T
9	105.5	16.7	138	4 AAM24182	Aam24182 Rhesus mo
10	105	16.6	140	4 AAB68882	Aab68882 Human REC
11	104	16.5	262	2 AAR97726	Aar97726 B10 singl
12	103	16.3	136	4 ABB11287	Abb11287 Human mem
13	103	16.3	329	6 ADB65555	Adb65555 Human pro
14	103	16.3	3931	6 ABU07377	Abu07377 Human pro
15	101.5	16.1	267	2 AAW04300	Aaw04300 Murine T-
16	99	15.7	246	5 AAE20273	Aae20273 Human lun
17	99	15.7	270	1 AAP50256	Aap50256 Sequence
18	98.5	15.6	252	5 ABP45484	Abp45484 Human Bly
19	98.5	15.6	740	4 ABG28146	Abg28146 Novel hum
20	97.5	15.4	216	5 ABP58185	Abp58185 CRIA-4-GC
21	96.5	15.3	125	5 ABP52181	Abp52181 Human mon
22	96.5	15.3	350	4 ABB69289	Abb69289 Drosophil
23	95.5	15.1	131	6 AAO39772	Aao39772 Rat myeli
24	95.5	15.1	150	2 AAW97817	Aaw97817 Rat myeli
25	94	14.9	526	2 AAW97814	Aaw97814 Human but

26	94	14.9	526	5 AAO15804	Aao15804 Human but
27	92.5	14.6	235	4 AAG64474	Aag64474 Human typ
28	92.5	14.6	388	5 ABB07681	Abb07681 MOG-Pc-fu
29	92.5	14.6	388	6 ADA14289	Ada14289 Mutated M
30	92.5	14.6	388	6 ADA14265	Ada14265 Human imm
31	92	14.6	174	3 AAB08208	Aab08208 Amino aci
32	91.5	14.5	223	2 AAY41083	Aay41083 Canine CT
33	91.5	14.5	268	2 AAR77288	Aar77288 T-cell re
34	91.5	14.5	269	3 AAY69995	Aay69995 Human rec
35	91	14.4	205	2 AAY41173	Aay41173 Llana Vhh
36	91	14.4	253	5 ABP45534	Abp45534 Human Bly
37	90.5	14.3	235	4 AAG64476	Aag64476 Human typ
38	90.5	14.3	502	6 ABJ37109	Abj37109 Concatame
39	89.5	14.2	110	6 ABR55818	Abr55818 Lambda ch
40	89.5	14.2	137	3 AAB08207	Aab08207 cDNA enco
41	89.5	14.2	383	5 ABB78103	Abb78103 Amino aci
42	89.5	14.2	383	5 AAU75121	Aau75121 Human sol
43	89.5	14.2	383	6 ABP56719	Abp56719 CTLA4 mut
44	89.5	14.2	383	7 ADD89007	Add89007 LI04EA29L
45	89.5	14.2	391	2 AAR97659	Aar97659 Single ch

## ALIGNMENTS

RESULT 1  
AAE02771  
ID AAE02771 standard; protein; 120 AA.  
XX  
AC AAE02771;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Human Nkp30 receptor extracellular region sequence.  
XX

Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;  
melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
therapy; extracellular region.  
XX

OS Homo sapiens.

PN WO200136630-A2.

XX 25-MAY-2001.

PD 15-NOV-2000; 2000MO-EP011697.

PR 15-NOV-1999; 99CA-02288307.

XX 15-NOV-1999; 99US-00440514.

PA (INNA-) INNATE PHARMA SAS.

XX (UUGE-) UNIV GENOVA.

PI Moretta A, Bottino C, Biassoni R;

XX WPI; 2001-329221/34.

Novel compound, useful for detection and/or quantifying the presence of  
NK cells, comprises the amino acid sequences of the Nkp30 molecule.

PS Claim 1; Fig 7B; 83pp; English.

XX The invention relates to human Nkp30 receptor and its corresponding cDNA  
molecule which is involved in natural cytotoxicity mediated by natural  
killer (NK) cells and antibodies that identify the same Nkp30 receptor  
is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively  
expressed on the surface of human mature NK cells. Nkp30 and its cDNA are  
useful for detecting and/or quantifying the presence of NK cells in a  
biological sample. The invention also provide kits for detecting and/or  
quantifying the presence of NK cells, for the selective removal of NK  
cells from a biological sample, for the positive and selective  
purification of NK cells from a biological sample and for the in vitro



stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (GVT) and especially graft versus leukaemia (GvL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. Nkp30 antibodies are useful for identifying Nkp30 natural ligands and allow assessment of the level of surface Nkp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence Nkp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the extracellular region of human Nkp30 receptor

XX Sequence 120 AA;

Query Match 100.0%; Score 632; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-62;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
 DB 1 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60

QY 61 PLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120  
 DB 61 PLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120

## RESULT 2

AY06402  
 ID AY06402 standard; protein; 177 AA.

XX AC AY06402;  
 XX 20-SEP-1999 (first entry)  
 XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
 XX BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
 XX signal transduction; immunomodulator; antiinflammatory;  
 XX autoimmune disease; inflammation; gene therapy; diagnosis.  
 XX Homo sapiens.

Key	Location/Qualifiers
FT Peptide	1..12
FT Protein	/note= "leader peptide"
FT Modified-site	13..177
FT Modified-site	/note= "mature protein"
FT Modified-site	42
FT Modified-site	/note= "N-glycosylated"
FT Modified-site	68
FT Modified-site	/note= "N-glycosylated"
FT Domain	139..162
FT Peptide	/note= "transmembrane domain"
FT Peptide	166..177
FT Peptide	/note= "alternatively spliced C-terminal end"

FN WO9923867-A2.

XX 20-MAY-1999.

XX 05-NOV-1998; 98WO-US023826.

XX 07-NOV-1997; 97US-0064761P.

XX (BIOJ ) BIOGEN INC.

XX Browning J;

XX WPI; 1999-418423/35.

DR N-PSDB; AAX59348.  
 XX Novel B-cell myelin oligodendrocyte glycoproteins.  
 XX Claim 2; Page 43; 43pp; English.

CC This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

XX Sequence 177 AA;

Query Match 100.0%; Score 632; DB 2; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-62;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 60  
 DB 19 LWSQPPEIRTLGSSAFPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGLA 78

QY 61 PLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 120  
 DB 79 PLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVGTGNGTRLVVEKEHPQLGAGT 138

## RESULT 3

AY06401  
 ID AAY06401 standard; protein; 190 AA.

XX AC AAY06401;

XX 20-SEP-1999 (first entry)

XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.

XX BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
 XX signal transduction; immunomodulator; antiinflammatory;  
 XX autoimmune disease; inflammation; gene therapy; diagnosis.

XX Homo sapiens.

Key	Location/Qualifiers
FT Peptide	1..12
FT Protein	/note= "leader peptide"
FT Modified-site	13..190
FT Modified-site	/note= "mature protein"
FT Modified-site	42
FT Modified-site	/note= "N-glycosylated"
FT Modified-site	68
FT Modified-site	/note= "N-glycosylated"
FT Modified-site	121
FT Domain	/note= "N-glycosylated"
FT Domain	139..162
FT Peptide	/note= "transmembrane domain"
FT Peptide	166..190
FT Peptide	/note= "alternatively spliced C-terminal end"

XX WO9923867-A2.

XX 20-MAY-1999.

XX 05-NOV-1998; 98WO-US023826.



XX 07-NOV-1997; 97US-0064761P.  
XX (BIOJ ) BIOGEN INC.  
XX Browning J;  
XX WPI; 1999-418423/35.  
XX N-PSDB; AAX59347.  
XX Novel B-cell myelin oligodendrocyte glycoproteins.  
XX Claim 2; Page 42; 43pp; English.  
XX This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)  
XX SQ Sequence 190 AA;  
Query Match 100.0%; Score 632; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 2.8e-62; Indels 0; Gaps 0;  
Matches 120; Conservative 0; Mismatches 0;  
Qy 1 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 60  
Db 19 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 78  
Qy 61 PLASSRFLHDHQAELHTRDVRGHDASIVYCVREVLGLGVGTGNGTRLVWEKEHPQLGAGT 120  
Db 79 PLASSRFLHDHQAELHTRDVRGHDASIVYCVREVLGLGVGTGNGTRLVWEKEHPQLGAGT 138  
RESULT 4  
AAE02769  
ID AAE02769 standard; protein; 190 AA.  
AC AAE02769;  
DT 06-AUG-2001 (first entry)  
XX Human Nkp30 receptor.  
DE Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;  
KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
KW therapy.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein /label= Signal\_peptide  
FT Region /label= Mature\_Nkp30\_receptor\_protein  
FT Region /label= Extracellular\_region  
FT Modified-site /note= "Forms an immunoglobulin (Ig) V-like domain"  
FT Modified-site /note= "N-glycosylation site"  
FT Modified-site 121  
FT Modified-site /note= "N-glycosylation site"

FT Region 139..157  
FT /label= Transmembrane\_region  
FT 158..190  
FT /label= Intracellular\_region  
XX WO200136630-A2.  
XX 25-MAY-2001.  
XX 15-NOV-2000; 2000WO-EP011697.  
XX 15-NOV-1999; 95CA-02288307.  
XX 15-NOV-1999; 95US-00440514.  
XX (INNA-) INNATE PHARMA SAS.  
XX (UYGE-) UNIV GENOVA.  
XX Moretta A, Bottino C, Biassoni R;  
XX WPI; 2001-329221/34.  
XX N-PSDB; AAD06564.  
XX Novel compound, useful for detection and/or quantifying the presence of  
XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.  
XX Claim 1; Fig 7B; 83pp; English.  
XX The invention relates to human Nkp30 receptor and its corresponding cDNA molecule which is involved in natural cytotoxicity mediated by natural killer (NK) cells and antibodies that identify the same. Nkp30 is selectively is a member of immunoglobulin super family (Ig-Sp). Nkp30 is selectively expressed on the surface of human mature NK cells. Nkp30 and its cDNA are useful for detecting and/or quantifying the presence of NK cells in a biological sample. The invention also provide kits for detecting and/or quantifying the presence of NK cells, for the selective removal of NK cells from a biological sample, for the positive and selective purification of NK cells from a biological sample and for the in vitro stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (Gvt) and especially graft versus leukaemia (GvL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. Nkp30 antibodies are useful for identifying Nkp30 natural ligands and allow assessment of the level of surface Nkp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence Nkp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is human Nkp30 receptor  
XX SQ Sequence 190 AA;  
Query Match 100.0%; Score 632; DB 4; Length 190;  
Best Local Similarity 100.0%; Pred. No. 2.8e-62; Indels 0; Gaps 0;  
Matches 120; Conservative 0; Mismatches 0;  
Qy 1 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 60  
Db 19 LWSQPEIRTELGSSAFPCSFNASQRLAIGSVTWFRDEVPGKEVRNGTPEFRGLA 78  
Qy 61 PLASSRFLHDHQAELHTRDVRGHDASIVYCVREVLGLGVGTGNGTRLVWEKEHPQLGAGT 120  
Db 79 PLASSRFLHDHQAELHTRDVRGHDASIVYCVREVLGLGVGTGNGTRLVWEKEHPQLGAGT 138  
RESULT 5  
AAY06403  
ID AAY06403 standard; protein; 201 AA.  
XX AAY06403;  
XX AAY06403;  
XX 20-SEP-1999 (first entry)  
XX

DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
 XX MBOG; B-cell myelin oligodendrocyte glycoprotein; human;  
 KW signal transduction; immunomodulator; antiinflammatory;  
 KW autoimmune disease; inflammation; gene therapy; diagnosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..12  
 FT Protein /note= "leader peptide"  
 FT Modified-site 13..201  
 FT Modified-site 42  
 FT Modified-site /note= "N-glycosylated"  
 FT Modified-site 68  
 FT Modified-site /note= "N-glycosylated"  
 FT Domain 121  
 FT Domain /note= "N-glycosylated"  
 FT Peptide 139..162  
 FT Peptide /note= "transmembrane domain"  
 FT Peptide 166..201  
 FT Peptide /note= "alternatively spliced C-terminal end"  
 XX  
 PN WO9923867-A2.  
 XX  
 PD 20-MAY-1999.  
 XX  
 PF 05-NOV-1998; 98WO-US023826.  
 XX  
 PR 07-NOV-1997; 97US-0064761P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Browning J;  
 XX  
 DR WPI; 1999-418423/35.  
 DR N-PSDB; AAX59349.  
 XX  
 PT Novel B-cell myelin oligodendrocyte glycoproteins.  
 PS Claim 2; Page 43; 43pp; English.  
 CC This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAX06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)  
 XX  
 SQ Sequence 201 AA;  
 Query Match 100.0%; Score 632; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 3e-62;  
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDVRVPGKEVRNGTPEFRGLA 60  
 DB 19 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDVRVPGKEVRNGTPEFRGLA 78  
 QY 61 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 120  
 DB 79 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 138

RESULT 6  
 AAE19109  
 ID AAE19109 standard; protein; 135 AA.  
 XX  
 AC AAE19109;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human Nkp30 protein.  
 XX  
 KW Human; natural killer cell activating protein; Nkp46; therapy; virucide;  
 KW viral infection; natural killer cell; NK; Nkp44; imaging agent; cancer;  
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; Nkp30.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200208287-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-IL000664.  
 XX  
 PR 20-JUL-2000; 2000IL-00137419.  
 XX  
 PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 PA (UTNE ) UNIV BEN-GURION NEGEV.  
 XX  
 PI Mandelboim O, Porgador A;  
 XX  
 DR WPI; 2002-195870/25.  
 DR N-PSDB; AAD30466.  
 XX  
 PT New targeting complex capable of targeting an active substance to a target cell, comprising a target recognition segment and an active PT segment, useful for treating pathologies associated with viral infections or cancer.  
 PT  
 PS Example 1; Page 108; 113pp; English.  
 XX  
 CC The invention relates to compositions and methods for the treatment and detection of a variety of viral infections, by using complex agents comprising the natural killer (NK) cells activating proteins, Nkp46 and Nkp44 and functional fragments thereof, linked to therapeutic or imaging agents. The complex is useful for treating pathologies associated with viral infections (e.g. infections caused by influenza virus, HIV, Epstein-Barr virus, cytomegalovirus, vaccinia virus, ECMV, MVM or herpes virus) and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for the imaging and monitoring of cancer. The complex may also be used to detect the presence of abnormal cells in a sample. The antibodies can be used to qualitatively or quantitatively detect the ligand for the complex. The present sequence is human Nkp30 protein  
 XX  
 SQ Sequence 135 AA;  
 Query Match 97.6%; Score 617; DB 5; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-61;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDVRVPGKEVRNGTPEFRGLA 60  
 DB 19 LWSQPPPIRTLEGSSAFPCSFNASQGRLAIGSVTFWFRDVRVPGKEVRNGTPEFRGLA 78  
 QY 61 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 117  
 DB 79 PLASSRFLHDHQAELHIRDVRGHDSIYVCVVEVLGLGVGTGNGTRLVVEKEHPQLG 135

RESULT 7  
 AAE19110  
 ID AAE19110 standard; protein; 369 AA.  
 XX  
 AC AAE19110;  
 XX

DT 29-AUG-2003 (revised)  
 DT 21-MAY-2002 (first entry)  
 XX Human Nkp30-IgG fusion protein.  
 DE Human, natural killer cell activating protein; Nkp46; therapy; virucide;  
 KW viral infection; natural killer cell; NK; Nkp44; imaging agent; cancer;  
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; IgG;  
 KW immunoglobulin G; fusion protein.  
 XX Homo sapiens.  
 OS Chimeric.  
 OS  
 FH Key Location/Qualifiers  
 FT Region 1. .135  
 FT /note= "Human Nkp30"  
 FT Region 136. .369  
 FT /note= "Human IgG"  
 XX WO200208287-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-IL000664.  
 XX  
 XX 20-JUL-2000; 2000IL-00137419.  
 XX  
 PA (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 PA (YINE ) UNIV BEN-GURION NEGEV.  
 XX  
 XX Mandelboim O, Porzador A;  
 XX  
 XX WPI; 2002-195870/25.  
 DR N-PSDB; AAD30467.  
 XX  
 XX New targeting complex capable of targeting an active substance to a  
 PT target cell, comprising a target recognition segment and an active  
 PT segment, useful for treating pathologies associated with viral infections  
 PT or cancer.  
 XX  
 XX Example 1; Page 108-110; 113pp; English.  
 PS  
 XX The invention relates to compositions and methods for the treatment and  
 CC detection of a variety of viral infections, by using complex agents  
 CC comprising the natural killer (NK) cells activating proteins, Nkp46 and  
 CC Nkp44 and functional fragments thereof, linked to therapeutic or imaging  
 CC agents. The complex is useful for treating pathologies associated with  
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein  
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECWV, MVM or herpes virus)  
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for  
 CC the imaging and monitoring of cancer. The complex may also be used to  
 CC detect the presence of abnormal cells in a sample. The antibodies can be  
 CC used to qualitatively or quantitatively detect the ligand for the  
 CC complex. The present sequence is human Nkp30- immunoglobulin G (IgG) Fc  
 CC region fusion protein. (Updated on 29-AUG-2003 to standardise OS field)  
 XX  
 SQ Sequence 369 AA;

Query Match 97.6%; Score 617; DB 5; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 3e-60;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LWSVQPEIRTLGGSAFLPCSNASQGRLAGSVTWFRDEVVPGKEVNGTPEFRGRILA 60  
 DB 19 LWSVQPEIRTLGGSAFLPCSNASQGRLAGSVTWFRDEVVPGKEVNGTPEFRGRILA 78  
 QY 61 PLASSRFLHDQAEHLIRDVGRGDASIIYVCRVEVLGLVGTGNGTGLVVEKEHPQLG 117  
 DB 79 PLASSRFLHDQAEHLIRDVGRGDASIIYVCRVEVLGLVGTGNGTGLVVEKEHPQLG 135

RESULT 8  
 AAU75565

ID AAU75565 standard; protein; 139 AA.  
 AC AAU75565;  
 XX  
 DT 23-APR-2002 (first entry)  
 DE Murine T cell receptor beta chain.  
 XX  
 KW Immunoglobulin superfamily; GP286; human; immunosuppressive;  
 KW autoimmune disease; rheumatoid arthritis; cancer; multiple sclerosis;  
 KW acquired immune deficiency syndrome; AIDS; inflammatory disorder;  
 KW pancreatic; antirheumatic; antipsoriatic; dermatologic; antianaemic;  
 KW cytostatic; antileukaemic; antiasthmatic; antiallergic;  
 KW T cell receptor beta chain.  
 XX  
 OS Mus sp.  
 XX  
 XX WO200200727-A2.  
 PN  
 XX 03-JAN-2002.  
 PD  
 XX 22-JUN-2001; 2001WO-US020038.  
 PF  
 XX 23-JUN-2000; 2000US-0213630P.  
 PR 13-APR-2001; 2001US-0283813P.  
 XX  
 XX (BIOJ ) BIOGEN INC.  
 PA  
 XX Carulli JP, Lukashin AV, Kilburn DR, Mathur P;  
 FI  
 XX WPI; 2002-090520/12.  
 DR  
 XX Isolated polynucleotide encoding a novel human immunoglobulin superfamily  
 PT member, named GP286, useful in the treatment of a disease condition that  
 PT relates to the immune system, e.g. a transplantation disorder or an  
 PT autoimmune disease.  
 PS  
 XX Disclosure; Page 167; 180pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide (I) encoding a novel  
 CC human immunoglobulin superfamily member, named GP286. The GP286  
 CC polynucleotide and polypeptide are useful for the treatment of a disease  
 CC condition that relates to the immune system, preferably a disease  
 CC condition that relates to T cells, e.g. a transplantation disorder, an  
 CC autoimmune disease (e.g. rheumatoid arthritis, systemic lupus  
 CC erythematosus, psoriasis, Sjogren's Syndrome, thyroiditis, Graves'  
 CC disease, pulmonary fibrosis, bronchiolitis obliterans, haemolytic anaemia  
 CC or Wegener's granulomatosis), cancer (e.g. leukaemia or lymphoma),  
 CC multiple sclerosis, graft versus host disease, Kawasaki syndrome,  
 CC acquired immune deficiency syndrome (e.g. AIDS) or an inflammatory  
 CC disorder (e.g. asthma, allergies, adult respiratory distress syndrome and  
 CC acute pancreatitis or chronic pancreatitis). The GP286 nucleic acids can  
 CC be used as probes for detecting, characterising and quantifying GP286  
 CC nucleic acids. The present sequence represents the amino acid sequence of  
 CC murine T cell receptor beta chain used in the method of the invention  
 XX  
 SQ Sequence 139 AA;

Query Match 18.8%; Score 118.5; DB 5; Length 139;  
 Best Local Similarity 30.2%; Pred. No. 4.6e-05;  
 Matches 39; Conservative 17; Mismatches 40; Indels 33; Gaps 7;

QY 2 WVS-----OPPRTLEGSSAFPCSNASQGRLAGSVTWFRDE-----VWPG 45  
 DB 18 WVSQDQVKQSPSALSLOEGTSSALRCNFS-----IATTTVQWFLQNSRGSNNLFLVLP- 71  
 QY 46 KEVENGPPEFRGRILAPLASSRFLHDQAEHLIRDVGRGDASIIYVCRVEVLGLG- VGTGN 103  
 DB 72 -----GTYE-NGRLLKSTFNK---ESYSTLHIDAQLEDSGTFCRAEVEGTGSKUSFGK 122  
 QY 104 GTRLVVEKE 112  
 DB 123 GAKLTVSPD 131

```
RESULT 9
AAM24182
ID AAM24182 standard; protein; 138 AA.
XX AC
XX DT 24-APR-2001 (first entry)
XX DE Human RECAP polypeptide, SEQ ID NO: 138.
XX DT 12-OCT-2001 (first entry)
XX DE Rhesus monkey EST encoded protein SEQ ID NO: 1707.
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
XX KW gene therapy; nutrition.
XX OS
XX PN Macaca mulatta.
XX DT WO200154477-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US002687.
XX PR 25-JAN-2000; 2000US-00491404.
XX PR 17-JUL-2000; 2000US-00617746.
XX PR 03-AUG-2000; 2000US-00631451.
XX PR 15-SEP-2000; 2000US-00663870.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX DT WPI; 2001-476164/51.
XX DT N-PSDB; AAH98841.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX PT antibodies and research use.
XX PS Claim 20; Page 1137-1138; 1275pp; English.
XX CC The present invention provides the protein and coding sequences of novel
XX CC proteins from a variety of organisms, including human, dog, cat, horse,
XX CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX CC from the organism of interest. They can be used in diagnostics,
XX CC forensics, gene mapping, identification of mutations, to assess
XX CC biodiversity, and for nutritional purposes. The present sequence is a
XX CC protein of the invention
XX SQ Sequence 138 AA;
Query Match 16.7%; Score 105.5; DB 4; Length 138;
Best Local Similarity 29.1%; Pred. No. 0.0013;
Matches 37; Conservative 19; Mismatches 46; Indels 25; Gaps 6;
QY 3 VSQPE-IRTEGSAFLPCSFNASQGRLAIGSVTFWRDEVPFGKEVRNGTPEFRGLAP 61
Db 22 VTQSPEALRLQEGESSLNGSYTVS-----GLRGLFWYRQD--PGK-----GPEFLFTLYS 70
QY 62 LASSR-----FLHDHQAELHIRDYRGHDASTYVCEVEVL-----GLGVGTGNGTRLV 108
Db 71 AGEKEKERLKAUTKKESFLHTAPEDSATYLCAVQAQFHSGGGADGLTFGKTRUK 130
QY 109 VEKEHPQ 115
Db 131 VLALYPE 137
RESULT 10
AAB68882
```

```
ID AAB68882 standard; protein; 140 AA.
XX AC
XX DT 24-APR-2001 (first entry)
XX DE Human RECAP polypeptide, SEQ ID NO: 12.
XX DT 12-OCT-2001 (first entry)
XX DE Human; RECAP; receptors and associated proteins; cerebroprotective;
XX KW norepinephrine; neuroprotective; anticonvulsant; antiparkinsonian; anti-HIV;
XX KW antidiabetic; immunostimulant; immunomodulator; antiinflammatory;
XX KW antithyroid; immunosuppressive; nephrotropic; antigout; thyromimetic;
XX KW cytosolic; antibacterial; virucide; fungicide; protozoicide;
XX KW antiarteriosclerotic; hepatotropic; gene therapy; infection; cancer.
XX OS Homo sapiens.
XX PN WO200107612-A2.
XX DT 01-FEB-2001.
XX PF 21-JUL-2000; 2000WO-US020035.
XX PR 21-JUL-1999; 99US-0145232P.
XX PR 07-OCT-1999; 99US-0158578P.
XX PR 12-NOV-1999; 99US-0165192P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Au-Young J, Bandman O, Tang YT, Yue H, Azimzai Y, Burford N;
XX PI Baughn MR, Lu DAM, Hillman JL, Patterson C, Lal P;
XX DT WPI; 2001-168554/17.
XX DT N-PSDB; AAF58606.
XX PT Novel receptors and associated proteins for diagnosis and treatment of
XX PT neurological disorders, immunological disorders including autoimmune/
XX PT inflammatory disorders and cell proliferative disorders such as cancer.
XX PS Claim 1; Page 104; 128pp; English.
XX CC The present sequence is a human RECAP (receptors and associated proteins)
XX CC polypeptide. RECAP polynucleotides and polypeptides are useful in the
XX CC diagnosis, treatment and prevention of neurological disorders such as
XX CC stroke, Alzheimer's disease, Pick's disease, Huntington's disease,
XX CC dementia, Parkinson's disease, Down's syndrome, amyotrophic lateral
XX CC sclerosis, multiple sclerosis, bacterial and viral meningitis, CJD
XX CC (Creutzfeldt-Jakob disease), GSS (Gerstmann-Strausler-Scheinker
XX CC syndrome), immunological disorders, including autoimmune/inflammatory
XX CC disorders such as AIDS, Digeorge's syndrome, severe combined
XX CC immunodeficiency disease (SCID), Chediak-Higashi syndrome, Cushing's
XX CC disease, Addison's disease, autoimmune thyroiditis, Crohn's disease,
XX CC diabetes mellitus, Good pasture's syndrome, gout, Grave's disease,
XX CC Hashimoto's thyroiditis, Sjogren's syndrome, Werner's syndrome, viral,
XX CC bacterial, fungal, parasitic, protozoal, and helminthic infections; and
XX CC cell proliferation disorders such as arteriosclerosis, atherosclerosis,
XX CC cirrhosis, hepatitis and cancer
XX SQ Sequence 140 AA;
Query Match 16.6%; Score 105; DB 4; Length 140;
Best Local Similarity 30.3%; Pred. No. 0.0015;
Matches 36; Conservative 20; Mismatches 37; Indels 26; Gaps 7;
QY 3 VSQPE-IRTEGSAFLPCSFNASQGRLAIGSVTFWRDEVPFGKEVRNGTPEFRGLAP 61
Db 25 VTQSPEALRLQEGESSLNGSYTVS-----GLRGLFWYRQD--PGK-----GPEFLFTLYS 73
QY 62 LASSR-----FLHDHQAELHIRDYRGHDASTYVCEVEVLGVGT-----GNGTRLV 109
Db 74 AGEKEKERLKAUTKKESFLHTAPEDSATYLCAVQD---GIGNVLHCGSGTQVV 129
```



CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, hematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a novel human  
 CC polypeptide of the invention  
 CC  
 CC Sequence 136 AA;

Query Match 16.3%; Score 103; DB 4; Length 136;  
 Best Local Similarity 28.7%; Pred. No. 0.0024;  
 Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;  
 QY 7 PEIRTEGSAFLPCSFNASQGLAIGSVTFWRDEVVGKVRNGTFEFGRLAPLASSR 66  
 Db 32 PDLSTTEGSHAFPCCKARGSP-----PNTWKD-----GQPVSGAEGK 71  
 QY 67 FLHDHQAEHLHIRDVRGHDASIYVCVE 93  
 Db 72 FTIQPSGELLVKNLEGQDAGTYTCTAE 98

RESULT 13  
 ADB63555  
 ID ADB63555 standard; protein; 329 AA.  
 AC ADB63555;  
 DT 04-DEC-2003 (first entry)  
 XX Human protein encoded by clone THYMU20028150.  
 DE Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;  
 KW cell regeneration; membrane protein; signal transduction-related protein;  
 KW transcription-related protein; osteoporosis; neurological disease;  
 KW cancer; tumour.  
 XX  
 OS Homo sapiens.  
 XX  
 FN EP1308459-A2.  
 XX  
 PD 07-MAY-2003.  
 XX  
 PF 28-MAR-2002; 2002EP-00007401.  
 XX  
 PR 05-NOV-2001; 2001JP-00379298.  
 XX  
 PR 25-JAN-2002; 2002US-00350978.  
 XX  
 PA (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX  
 DR WPI; 2003-450961/43.  
 DR N-PSDB; ADB63585.  
 XX

PT New polynucleotides and polypeptides, useful for developing a diagnostic  
 PT marker or medicines for regulation of their expression and activity, or

PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

XX  
 CC The invention discloses a polynucleotide comprising a sequence selected  
 CC from 1970 fully defined nucleotide sequences which encode novel  
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide  
 CC or its partial peptide, an antibody binding to the polypeptide or peptide  
 CC of the polynucleotide, immunologically assaying the polypeptide or  
 CC peptide of the polynucleotide by contacting the polypeptide or peptide  
 CC with the antibody of the encoded protein, and observing the binding  
 CC between the two, a transformant carrying the polynucleotide in an  
 CC expressible manner and an antisense polynucleotide. The oligonucleotide  
 CC is useful as a primer for synthesising the polynucleotide, or as a probe  
 CC for detecting the polynucleotide. The polynucleotides and encoded  
 CC proteins are useful as pharmaceutical agents and many disease-related  
 CC genes may be included in them for developing a diagnostic marker or  
 CC medicines for regulation of their expression and activity, or as targets  
 CC of gene therapy. The genes are involved in tissue and/or cell  
 CC regeneration. Membrane proteins, signal transduction-related proteins,  
 CC transcription-related proteins, disease-related proteins and genes  
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,  
 CC neurological diseases, cancer, tumours). The cDNA may be used to regulate  
 CC the activity or expression of the encoded protein to treat diseases. The  
 CC sequence presented is a protein of the invention. Note: Some of the  
 CC sequence data for this patent is not represented in the printed  
 CC specification, but is based on sequence information supplied by the  
 CC European Patent Office.  
 CC  
 CC Sequence 329 AA;

Query Match 16.3%; Score 103; DB 7; Length 329;  
 Best Local Similarity 28.7%; Pred. No. 0.0074;  
 Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;  
 QY 7 PEIRTEGSAFLPCSFNASQGLAIGSVTFWRDEVVGKVRNGTFEFGRLAPLASSR 66  
 Db 176 PDLSTTEGSHAFPCCKARGSP-----PNTWKD-----GQPVSGAEGK 215  
 QY 67 FLHDHQAEHLHIRDVRGHDASIYVCVE 93  
 Db 216 FTIQPSGELLVKNLEGQDAGTYTCTAE 242

RESULT 14  
 ADBU07377  
 ID ADBU07377 standard; protein; 3931 AA.  
 XX  
 AC ADBU07377;  
 XX  
 DT 28-JAN-2003 (first entry)  
 XX  
 DE Human protein NOV9.  
 XX  
 KW Human; NOV9; cardiomyopathy; atherosclerosis; cancer; hypertension;  
 KW diabetes; inflammation; autoimmune disorder; allergy; blood disorder;  
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;  
 KW immunoglobulin A nephropathy; cirrhosis; arthritis; Alzheimer's disease;  
 KW Parkinson's disease; goitre; infection; stroke; muscular dystrophy;  
 KW epilepsy; wasting disorder; neurogenesis; cell differentiation;  
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;  
 KW Gene therapy; single nucleotide polymorphism; SNP.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200285922-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 11-APR-2002; 2002WO-US011634.  
 XX  
 PR 23-APR-2001; 2001US-0285748P.  
 PR 24-APR-2001; 2001US-0286068P.





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 41.6976 Seconds  
(without alignments)  
607.670 Million cell updates/sec

Title: US-10-036-444-4

Perfect score: 632  
Sequence: 1 LWVSQPEIRTEGSSAFPLP.....TGNTRLVVEKHPOLGAGT 120

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pcp:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pcp:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pcp:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pcp:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pcp:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pcp:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pcp:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pcp:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pcp:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pcp:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pcp:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pcp:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pcp:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pcp:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pcp:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	632	100.0	120	13	US-10-036-444-4
2	632	100.0	190	13	Sequence 4, Appli
3	118.5	18.8	139	14	Sequence 2, Appli
4	103	16.3	329	15	Sequence 6, Appli
5	103	16.3	329	15	Sequence 3709, Ap
6	99	15.7	246	10	Sequence 18, Appli
7	98.5	15.6	252	10	Sequence 49, Appli
8	94	14.9	526	9	Sequence 1495, Ap
9	94	14.9	526	9	Sequence 19, Appli
10	94	14.9	526	9	Sequence 13, Appli
11	94	14.9	526	14	Sequence 19, Appli
12	93	14.7	1598	10	Sequence 7, Appli
13	93	14.7	1694	10	Sequence 59, Appli
14	91	14.4	205	9	Sequence 57, Appli
15	91	14.4	253	10	Sequence 9, Appli
					Sequence 1545, Ap

16	90.5	14.3	502	14	US-10-363-427-24	Sequence 24, Appli
17	89.5	14.2	110	14	US-10-269-805-50	Sequence 50, Appli
18	89.5	14.2	383	10	US-09-898-195A-11	Sequence 11, Appli
19	89.5	14.2	383	14	US-10-057-288-6	Sequence 6, Appli
20	89.5	14.2	383	14	US-10-155-514-10	Sequence 10, Appli
21	89.5	14.2	383	16	US-10-419-008-11	Sequence 11, Appli
22	88.5	14.0	223	9	US-09-989-545-20	Sequence 20, Appli
23	88.5	14.0	223	14	US-10-211-207-5	Sequence 5, Appli
24	88.5	14.0	223	14	US-10-077-106-5	Sequence 5, Appli
25	88	13.9	226	13	US-10-038-107A-1	Sequence 1, Appli
26	88	13.9	226	15	US-10-371-069-32	Sequence 32, Appli
27	88	13.9	226	15	US-10-371-645-32	Sequence 32, Appli
28	88	13.9	226	15	US-10-371-260-32	Sequence 32, Appli
29	88	13.9	226	15	US-10-411-010-19	Sequence 19, Appli
30	88	13.9	524	14	US-10-156-424A-9	Sequence 9, Appli
31	87.5	13.8	151	14	US-10-029-386-32704	Sequence 32704, A
32	87.5	13.8	151	14	US-10-207-655-309	Sequence 309, App
33	87.5	13.8	212	10	US-09-898-195A-17	Sequence 17, Appli
34	87.5	13.8	212	14	US-10-057-288-12	Sequence 12, Appli
35	87.5	13.8	212	14	US-10-155-514-2	Sequence 2, Appli
36	87.5	13.8	212	16	US-10-419-008-17	Sequence 17, Appli
37	87.5	13.8	247	9	US-09-893-348-14	Sequence 14, Appli
38	87.5	13.8	247	10	US-09-218-277-14	Sequence 14, Appli
39	87.5	13.8	247	15	US-10-245-871-877	Sequence 877, App
40	87.5	13.8	374	13	US-10-027-075-26	Sequence 26, Appli
41	87.5	13.8	374	13	US-10-027-075-28	Sequence 28, Appli
42	87.5	13.8	377	13	US-10-027-075-24	Sequence 24, Appli
43	87.5	13.8	377	14	US-10-363-427-16	Sequence 16, Appli
44	87.5	13.8	382	14	US-10-207-655-307	Sequence 307, App
45	87.5	13.8	382	14	US-10-207-655-316	Sequence 316, App

ALIGNMENTS

RESULT 1

US-10-036-444-4  
; Sequence 4, Application US/10036444  
; Publication No. US20020142445A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; APPLICANT: UNIVERSITA DI GENOVA  
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and  
; TITLE OF INVENTION: antibodies that identify the same"  
; FILE REFERENCE: SEQ-PR-1060  
; CURRENT APPLICATION NUMBER: US/10/036.444  
; PRIORITY FILING DATE: 2002-01-07  
; PRIORITY FILING DATE: 09/440,514  
; PRIORITY FILING DATE: 1999-11-15  
; PRIORITY FILING DATE: 1999-12-07  
; PRIORITY FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Human NK cell  
US-10-036-444-4

Query Match	100.0%;	Score 632;	DB 13;	Length 120;			
Best Local Similarity	100.0%;	Pred. No. 8e-62;	Mismatches 0;	Indels 0; Gaps 0;			
Matches 120;	Conservative 0;						
QY	1	LWVSQPEIRTEGSSAFPLP	CSFNASQGR	LAIGSVTWFR	DEVVPGKEVRN	GTPEFRGLA	60
Db	1	LWVSQPEIRTEGSSAFPLP	CSFNASQGR	LAIGSVTWFR	DEVVPGKEVRN	GTPEFRGLA	60
QY	61	PLASRFLHDQAEHLH	VRGHDASIV	VCVEVLGL	VGVTGNGTRL	WVEKEH	POLGAGT 120
Db	61	PLASRFLHDQAEHLH	VRGHDASIV	VCVEVLGL	VGVTGNGTRL	WVEKEH	POLGAGT 120



```
Query Match      16.3%; Score 103; DB 15; Length 329;
Best Local Similarity 28.7%; Fred.No. 0.005;
Matches 25; Conservative 11; Mismatches 31; Indels 20; Gaps 2;

Qy 7 PEIITLEGSAFLPCSFNASQGRLAIGSVTFWFRDEVVVGKEVRNGTDEFGRIAPLASRR 66
    :::::::::::
Db 176 PDLSTTEGSHAFLEFCARGSPF-----PNIITWDKD-----QQPVSGAGBK 215
                               :
Qy 67 FLIDHQAEELHIRDVVRGHDAIIYVCERVE 93
Db 216 FTTPSGELLVKNLGGODAGTYICTAAE 242
```

RESULT 5  
US-10-120-801-18  
; Sequence 18, Application US/10120801  
; Publication No. US20030203843A1  
; GENERAL INFORMATION:  
; APPLICANT: Pena, Carol  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Mehraban, Fuad  
; APPLICANT: Topper, James N.  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Wasserman, Scott  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Smithson, Glennida  
; APPLICANT: Gunther, Erik  
; APPLICANT: Komuves, Laszlo  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-340  
; CURRENT APPLICATION NUMBER: US/10/120,801  
; CURRENT FILING DATE: 2002-04-11  
; PRIOR APPLICATION NUMBER: 60/2885748  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: 60/286068  
; PRIOR FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: 60/286292  
; PRIOR FILING DATE: 2001-04-25  
; PRIOR APPLICATION NUMBER: 60/288334  
; PRIOR FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 60/291241  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 60/322284  
; PRIOR FILING DATE: 2001-09-14  
; PRIOR APPLICATION NUMBER: 60/285609



; APPLICANT: Yoshinaga, Steven Kiyoshi  
; TITLE OF INVENTION: B7-Like Polypeptides and Uses Thereof  
; FILE REFERENCE: 00-759-A  
; CURRENT APPLICATION NUMBER: US/09/955,866  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/233,867  
; PRIOR FILING DATE: 2000-09-20  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-955-866-13

Query Match 14.9%; Score 94; DB 9; Length 526;  
Best Local Similarity 39.6%; Pred. No. 0.088;  
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;  
QY 6 PPE-IRTEGSSAFPC--SFNASQGLAIGSVTFWFRDEVVPGKEV-RNG-----TPE 54  
DB 34 PPEPLAVVGDAELPCRLSPNASEHL---ELRFRKVKVSPAVLVHRDGRGEQAEQMP 90  
QY 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDSIYVC 90  
DB 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDDG EYTC 124

RESULT 10  
US-09-896-738-19  
; Sequence 19, Application US/09896738  
; Patent No. US20020165347A1  
; GENERAL INFORMATION:  
; APPLICANT: Fox, Michael  
; APPLICANT: Sullivan, John K.  
; TITLE OF INVENTION: B7-Like Molecules and Uses Thereof  
; FILE REFERENCE: 00-513-A  
; CURRENT APPLICATION NUMBER: US/09/896,738  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 60/215,645  
; PRIOR FILING DATE: 2000-06-30  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-896-738-19

Query Match 14.9%; Score 94; DB 9; Length 526;  
Best Local Similarity 39.6%; Pred. No. 0.088;  
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;  
QY 6 PPE-IRTEGSSAFPC--SFNASQGLAIGSVTFWFRDEVVPGKEV-RNG-----TPE 54  
DB 34 PPEPLAVVGDAELPCRLSPNASEHL---ELRFRKVKVSPAVLVHRDGRGEQAEQMP 90  
QY 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDSIYVC 90  
DB 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDDG EYTC 124

RESULT 11  
US-10-156-424A-7  
; Sequence 7, Application US/10156424A  
; Publication No. US20030039999A1  
; GENERAL INFORMATION:  
; APPLICANT: Yoshinaga, Steve Kiyoshi  
; APPLICANT: Suh, Woong-Kyung  
; APPLICANT: Mak, Tak W.  
; TITLE OF INVENTION: B7 Related Protein-2 Molecules and Uses Thereof  
; FILE REFERENCE: 01-384-A

; CURRENT APPLICATION NUMBER: US/10/156,424A  
; CURRENT FILING DATE: 2002-09-10  
; PRIOR APPLICATION NUMBER: 60/293,629  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-156-424A-7

Query Match 14.9%; Score 94; DB 14; Length 526;  
Best Local Similarity 39.6%; Pred. No. 0.088;  
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;  
QY 6 PPE-IRTEGSSAFPC--SFNASQGLAIGSVTFWFRDEVVPGKEV-RNG-----TPE 54  
DB 34 PPEPLAVVGDAELPCRLSPNASEHL---ELRFRKVKVSPAVLVHRDGRGEQAEQMP 90  
QY 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDSIYVC 90  
DB 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDDG EYTC 124

RESULT 12  
US-09-863-776-59  
; Sequence 59, Application US/09863776  
; Publication No. US20030198953A1  
; GENERAL INFORMATION:  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Tchernev, Vellizar T  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Shenoy, Suresh G  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Li, Li  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Gangolli, Esha  
; TITLE OF INVENTION: No. US20030198953A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-020  
; CURRENT APPLICATION NUMBER: US/09/863,776  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: 09/540,763  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: 60/206,679  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 60/206,688  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 60/206,829  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 60/207,748  
; PRIOR FILING DATE: 2000-05-30  
; PRIOR APPLICATION NUMBER: 60/207,798  
; PRIOR FILING DATE: 2000-05-30  
; PRIOR APPLICATION NUMBER: 60/208,263  
; PRIOR FILING DATE: 2000-05-31  
; PRIOR APPLICATION NUMBER: 60/208,831  
; PRIOR FILING DATE: 2000-06-02  
; PRIOR APPLICATION NUMBER: 60/209,451  
; PRIOR FILING DATE: 2000-06-05  
; PRIOR APPLICATION NUMBER: 60/210,060  
; PRIOR FILING DATE: 2000-06-07  
; PRIOR APPLICATION NUMBER: 60/219,507  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/221,337  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/221,927  
; PRIOR FILING DATE: 2000-07-31  
; PRIOR APPLICATION NUMBER: 60/263,135  
; PRIOR FILING DATE: 2001-01-19

```

, , PRIOR APPLICATION NUMBER: 60/263,688
, , PRIOR FILING DATE: 2001-01-24
, , PRIOR APPLICATION NUMBER: 60/263,694
, , PRIOR FILING DATE: 2001-01-24
, , NUMBER OF SEQ ID NOS: 155
, , SOFTWARE: Patent.in Ver. 2.1
, , SEQ ID NO 59
, , LENGTH: 1598
, , TYPE: prt
, , ORGANISM: Mus musculus
US-09-863-776-59

```

```

Query Match      14.7%; Score 93; DB 10; Length 1598;
Best Local Similarity 29.4%; Pred. No. 0.45;
Matches 32; Conservative 17; Mismatches 34; Indels 26; Gaps 8;

QY      2 W-VOPPIRTLEGGSAFLPCSFN-----ASGRLAIGSVTFRDEVVPQKE---VRNG 51
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      21 WGVSPKQVQGLSGCLLPCIFSPADVPVSGITAI---WYD--YSGRQVVIHSG 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      52 TPE-----FRGLAPLASSRFLDHQ-AELHIRDVAGHDASIVYCRVEV 94
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      75 DPKLVDFKFRGRGAELGMNN---DHKCNLLLLKDLKPEDSGTVNPFBI 119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 13
US-09-863-576-57
US-Sequence 576, Application US/09863776
Publication No. US20030198953A1
GENERAL INFORMATION:
APPLICANT: Spytek, Kimberly A
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T
APPLICANT: Mishra, Vishnu
APPLICANT: Padigaru, Muralidhara
APPLICANT: Spaderna, Steven K
APPLICANT: Shenoy, Suresh G
APPLICANT: Raastelli, Luca
APPLICANT: Li, Li
APPLICANT: Taupier, Raymond J
APPLICANT: Gangolli, Esha
TITLE OF INVENTION: No. US20030198953
FILE REFERENCE: 21402-020
CURRENT APPLICATION NUMBER: US/09/863-576
CURRENT FILING DATE: 2003-05-23
PRIOR APPLICATION NUMBER: 09/540,766
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/206,671
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: 60/206,681
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/206,821
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/207,741
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: 60/207,791
PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: 60/208,261
PRIOR FILING DATE: 2000-05-31
PRIOR APPLICATION NUMBER: 60/208,831
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: 60/209,451
PRIOR FILING DATE: 2000-06-05
PRIOR APPLICATION NUMBER: 60/210,061
PRIOR FILING DATE: 2000-06-07
PRIOR APPLICATION NUMBER: 60/219,501
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: 60/221,331
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/221,921
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: 60/263,131
PRIOR FILING DATE: 2003-01-19

```

```

; PRIOR APPLICATION NUMBER: 60/263,688
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/263,694
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 57
; LENGTH: 1694
; TYPE: prt
; ORGANISM: Mus musculus
US-09-863-776-57

```

```

Query Match      14.7%; Score 93; DB 10; Length 1694;
Best Local Similarity 29.4%; Pred. No. 0.48; 34; Indels 26; Gaps 8;
Matches 3; Conservative 17; Mismatches 17;

Qy 2 W-VSQPPEIHTLGGSAFTPCSN-----ASGRLAIGSVTFRDEVDPKE---VRNG 51
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 21 WGVSSPKNVGGSCLLPFCISYPADVPVNSGITAL-----WYD--YSGKQVVIHS 74
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy 52 TPE-----FRGLAPLASRFLDHQ-AELHIRDVGRGDASIYVCRVVEV 94
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 75 DPKLVDRFRGAELMGNM-----DHKVCNLLLKDLKPEDSGTNYFAFEI 119
      :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

RESULT 14
US-09-252-150-9
; Sequence 9, Application US/09252150A
; Patent No. US20020155604A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden Ledbetter, Martha
; APPLICANT: Brady, William A.
; APPLICANT: Grosmaire, Laura S.
; APPLICANT: Law, Che-Leung
; APPLICANT: Dua, Raj
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATING
; FILE REFERENCE: 9113-0019-999
; CURRENT APPLICATION NUMBER: US/09/252,150A
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US 60/075,274
; EARLIER FILING DATE: 1998-02-19
; EARLIER APPLICATION NUMBER: US 60/108,683
; EARLIER FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Llama llama
US-09-252-150-9

```

Query Match	14.4%	Score 91;	DB 9;	Length 205;
Best Local Similarity	28.0%	Pred. No. 0.058;		
Matches 35; Conservative 14;	Mismatches 36;	Indels 40;	Gaps 8;	
Qy	14	GSSAFPCSFNASQGR---	LAIGSVTFWDFEVVPGKEVNGTPEPRGRRLAPLASSFLH	69
Db	15	GDSLRSLC---AASGRFTNYAMG---	WFRQ---APGKE-----PEFVARISRVGSSTFYT	61
Qy	70	D-----HQAELHIRDVRGHIDASIIYVCRVEVLGLGVGT-----	GNGRFL	107
Db	62	DSVKGRFTISRDNAKNTMWLQMSMKAEADTAVVYCAAD-	SDYGPGRSSEYDYGQGTQV	120
Qy	108	VVEKE	112	
Db	121	TVSSE	125	

RESULT 15  
US-09-880-748-1545  
; Sequence 1545, Application US/09880748

```

; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PFS23
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: Patent in ver. 2.0
; SEQ ID NO 1545
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1545

Query Match      14.4%; Score 91; DB 10; Length 253;
Best Local Similarity 25.0%; Pred. No. 0.076;
Matches 31; Conservative 20; Mismatches 37; Indels 36; Gaps 6;

QY      3 VSQPEIRTLGSSAFPCSFNASQGRLAGSVTWPRD-----EVVPGKEVRNGTPEF 55
Db      147 LTQPPSVSAASQQRVTIPCS--GSSNIGIGSVSWYQPFSGAPKFLISGNKPSGIPD- 203
QY      56 RGRAPLASSRFLHDHQAELHIRDVRGHDAIYVCRVEVLGLGVGT-----GNGT 105
Db      204 RFSYKSGTS-----ATLDTGLQTGDEADYYC-----GTWDTTINAWVFGAGT 247
QY      106 RLIV 109
Db      248 KLTV 251

```

Search completed: February 26, 2004, 12:33:53  
Job time : 43.6976 secs



Db 71 PGTKE-NGRKSTNSK---ESYSTLHRAQLEDSTYFCAAE-----GGGSNYKLTFGK 122  
Qy 104 GTRLVW 109  
Db 123 GTLLTV 128

RESULT 2  
US-09-082-593-10  
; Sequence 10, Application US/09082593  
; Patent No. 6180104  
; GENERAL INFORMATION:  
; APPLICANT: DAVIS, MARK M.  
; APPLICANT: HEDRICK, STEPHEN M.  
; TITLE OF INVENTION: T CELL RECEPTOR BETA SUBUNIT  
; FILE REFERENCE: JKI193-195DIV2  
; CURRENT APPLICATION NUMBER: US/09/082,593  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 270  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-082-593-10

Query Match 15.7%; Score 99; DB 3; Length 270;  
Best Local Similarity 29.4%; Pred. No. 0.0018;  
Matches 37; Conservative 17; Mismatches 42; Indels 30; Gaps 8;  
Qy 3 VSOPPE-IRTEGSSAFPCSFNASQGLAIGSVTWFRDEVPVQKEVR-----NGTPE 54  
Db 24 VQSPESILVPEGARTSLNCTFSDASQY-----FWYRQH--SGKAPALMSIFSNGEKE 77  
Qy 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDASIYVCRVEVLGLGVTGNGTFLVVEKEHP 114  
Db 78 -EGRTIHLNKASLH---PSLHSDSPSDSALYLCATVLYG-----GSGNKLI----- 122  
Qy 115 QLQAGT 120  
Db 123 -FGTGT 127

RESULT 3  
US-10-026-045-1  
; Sequence 1, Application US/10026045  
; Patent No. 6573236  
; GENERAL INFORMATION:  
; APPLICANT: Genain, Claude P  
; APPLICANT: Hauser, Stephen L  
; TITLE OF INVENTION: Inhibiting MOG-Antibody Binding  
; FILE REFERENCE: UCSF99-020-3  
; CURRENT APPLICATION NUMBER: US/10/026,045  
; CURRENT FILING DATE: 2003-01-13  
; PRIOR APPLICATION NUMBER: US 09/384,036  
; PRIOR FILING DATE: 1999-08-26  
; PRIOR APPLICATION NUMBER: us 60/097,953  
; PRIOR FILING DATE: 1998-08-26  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 218  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-026-045-1  
Query Match 15.1%; Score 95.5; DB 4; Length 218;  
Best Local Similarity 36.2%; Pred. No. 0.0035;  
Matches 34; Conservative 8; Mismatches 33; Indels 19; Gaps 6;  
Qy 9 IRTLEGSSAFPCSFNASQGLAIG--SVTWFRDEVPVQKEV-----RNG-----TPFR 56

Db 12 IRALVGDAAELPC--RISPGKNATMEVGWYRS---PFSRVVHLRYRNGKQDAEQAPPEYR 66  
Qy 57 GRPLAPLASSRFLHDHQAELHIRDVRGHDASIYVC 90  
Db 67 GRTELLKES--IGEGKVALRIQNVRFSDGGYTC 98

RESULT 4  
US-09-910-174B-9  
; Sequence 9, Application US/09910174B  
; Patent No. 6630575  
; GENERAL INFORMATION:  
; APPLICANT: Coyle, Anthony J.  
; APPLICANT: Fraser, Christopher C.  
; APPLICANT: Manning, Stephen  
; TITLE OF INVENTION: B7-H2 Molecules, No. 6630575el Members of the B7  
; TITLE OF INVENTION: Family and Uses Thereof  
; FILE REFERENCE: 35800/236924  
; CURRENT APPLICATION NUMBER: US/09/910,174B  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 09/620,461  
; PRIOR FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-910-174B-9

Query Match 14.9%; Score 94; DB 4; Length 526;  
Best Local Similarity 39.6%; Pred. No. 0.017;  
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;  
Qy 6 PPE-IRTEGSSAFPC--SFNASQGLAIGSVTWFRDEVPVQKEV-RNG-----TPE 54  
Db 34 PPEPLAVVGEDAEFLCRLSPNASAEHL---ELRFRKKVSPAVLVHRDGRQEAQMP 90  
Qy 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDASIYVC 90  
Db 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDDGEYTC 124

RESULT 5  
US-09-620-461-9  
; Sequence 9, Application US/09620461  
; Patent No. 6635750  
; GENERAL INFORMATION:  
; APPLICANT: Coyle, Anthony J.  
; APPLICANT: Fraser, Christopher C.  
; APPLICANT: Manning, Stephen  
; TITLE OF INVENTION: B7-H2 Molecules, No. 6635750el Members of the B7  
; TITLE OF INVENTION: Family and Uses Thereof  
; FILE REFERENCE: 5800-149  
; CURRENT APPLICATION NUMBER: US/09/620,461  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 9  
; LENGTH: 526  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-620-461-9

Query Match 14.9%; Score 94; DB 4; Length 526;  
Best Local Similarity 39.6%; Pred. No. 0.017;  
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;  
Qy 6 PPE-IRTEGSSAFPC--SFNASQGLAIGSVTWFRDEVPVQKEV-RNG-----TPE 54  
Db 34 PPEPLAVVGEDAEFLCRLSPNASAEHL---ELRFRKKVSPAVLVHRDGRQEAQMP 90  
Qy 55 FRGLAPLASSRFLHDHQAELHIRDVRGHDASIYVC 90

Db 91 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDGGEYTC 124

## RESULT 6

US-08-724-394A-1  
; Sequence 1, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Krommal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchinashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724.394A  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 589 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Region  
; LOCATION: 1..589  
; OTHER INFORMATION: /note= "BT"

US-08-724-394A-1

Query Match 14.9%; Score 94; DB 2; Length 589;  
Best Local Similarity 39.6%; Pred. No. 0.019;  
Matches 38; Conservative 5; Mismatches 37; Indels 16; Gaps 7;

QY 6 PPE-IRTEGSAFLPC--SFNASQRLAIGSVTWFRDEVVPGKEV-RNG-----TPE 54  
Db 37 PPEPLAVVGDEALPCRLSPNASAEHL---ELRFRKKVSPAVLVHRDGRQEAQMP 93

QY 55 FGRLAPLASSRFLHDHQAELHIRDVGRGDASIYVC 90  
Db 94 YRGR-ATLVQDGIAGRVA-LRIRGVRVSDGGEYTC 127

## RESULT 7

US-08-804-180C-4

; Sequence 4, Application US/08804180C

; Patent No. 6107056

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

APPLICANT: Martin K. Oaks  
TITLE OF INVENTION: SCTL-4 and Its Soluble Products  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Thomas M. Wozny  
STREET: 100 East Wisconsin Avenue  
CITY: Milwaukee  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.50 inch Disk  
COMPUTER: IBM  
OPERATING SYSTEM: DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804.180C  
FILING DATE: February 20, 1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas M. Wozny  
REGISTRATION NUMBER: 28,922  
REFERENCE/DOCKET NUMBER: 3284-00003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (414) 271-7590  
TELEFAX: (414) 271-5770  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: Polypeptide  
HYPOTHETICAL: no  
FRAGMENT TYPE: Internal fragment  
ORIGINAL SOURCE:  
ORGANISM: Rattus No. 6107056vegicus  
STRAIN: ACI  
DEVELOPMENTAL STAGE: Adult  
CELL TYPE: Splenocyte  
FEATURE:  
NAME/KEY: Rat SCTL-4  
IDENTIFICATION METHOD: Found by experiment  
OTHER INFORMATION: B7 binding protein  
US-08-804-180C-4

Query Match 14.6%; Score 92; DB 3; Length 174;  
Best Local Similarity 29.5%; Pred. No. 0.0067;  
Matches 36; Conservative 19; Mismatches 49; Indels 18; Gaps 7;

QY 3 VSQPPFIRTEGSSAFPC----SFNASQRLAIGSVTWFRDEVVPGKEV-RNGTPEFRGR 58  
Db 40 VTQPSVVLASSHGVSFPCFVYASHNTDEVVTVLROT--NDQVT---EVCAITFTVKNT 94

QY 59 LA----PLASSRFLHDHQAELHIRDVGRGDASIYVCVEVL--GLGVGTGNGTGL-VVE 110  
Db 95 LQFLDDPFCGTF--NESRVNLTQGLRAADTGLYFCKVELMYPFVFGMNGTQIYVIA 153

QY 111 KE 112  
Db 154 KE 155

RESULT 8  
US-09-171-945-89  
; Sequence 89, Application US/09171945  
; Patent No. 6277599  
; GENERAL INFORMATION:  
; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek

; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said

; TITLE OF INVENTION: Antibody, and Their Therapeutic Use in an Adept System

; GENERAL INFORMATION:

; APPLICANT: Emery, Stephen

; APPLICANT: Copley, Clive Graham

; APPLICANT: Edge, Michael Derek



FILE REFERENCE: Monoclonal Antibody to CEA  
CURRENT APPLICATION NUMBER: US/09/171,945  
CURRENT FILING DATE: 1998-10-29  
PRIOR APPLICATION NUMBER: GB9703103.3  
PRIOR FILING DATE: 1997-02-14  
PRIOR APPLICATION NUMBER: GB9609405.7  
PRIOR FILING DATE: 1996-05-04  
PRIOR APPLICATION NUMBER: PCT/GB97/01165  
PRIOR FILING DATE: 1997-04-29  
NUMBER OF SEQ ID NOS: 131  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 89  
LENGTH: 120  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: humanized

US-09-171-945-89

Query Match 14.2%; Score 99.5; DB 3; Length 120;  
Best Local Similarity 27.8%; Pred. No. 0.008;  
Matches 32; Conservative 10; Mismatches 34; Indels 39; Gaps 6;

QY 2 WVSOPPEIRTELGSSAFPCSFNASQGRLAIGSVTWRDEVPKGVRNG---TPFRG 57  
DB 36 WVRQPP-----GR-GLGWIGWIDPE-----NGDTEYAPKFRG 66

QY 58 RLAPASSRFLHQALHIRDVGHDSIVYCVREVLGLGVGT---GNGTRLVV 109  
DB 67 RATMLADS---SKNQASLESLSVTAADTAIVYCHVLIYAGYAMDYWGQGLTVV 118

#### RESULT 9

US-08-804-180C-2

Sequence 2, Application US/08804180C  
Patent No. 6107056

GENERAL INFORMATION:

APPLICANT: Martin K. Oaks

TITLE OF INVENTION: sCTLA-4 and Its Soluble Products

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSES: Thomas M. Wozny

STREET: 100 East Wisconsin Avenue

CITY: Milwaukee

STATE: Wisconsin

COUNTRY: USA

ZIP: 53202

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.50 inch Disk

COMPUTER: IBM

OPERATING SYSTEM: DOS

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,180C

FILING DATE: February 20, 1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Thomas M. Wozny

REGISTRATION NUMBER: 28,922

REFERENCE/DOCKET NUMBER: 3284-00003

TELEPHONE: (414) 271-7590

TELEFAX: (414) 271-5770

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 137

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

DESCRIPTION: no

HYPOTHETICAL: no

FRAGMENT TYPE: Mature Polypeptide

ORIGINAL SOURCE:

ORGANISM: Homo Sapien  
DEVELOPMENTAL STAGE: Adult  
TISSUE TYPE: Lymphnode  
FEATURE:  
NAME/KEY: Human sCTLA-4  
IDENTIFICATION METHOD: Found by experiment  
OTHER INFORMATION: Asn 76 and Asn 108 represent N-linked glycosylation; B7 bind  
US-08-804-180C-2

Query Match 14.2%; Score 89.5; DB 3; Length 137;

Best Local Similarity 25.2%; Pred. No. 0.0095;

Matches 33; Conservative 18; Mismatches 59; Indels 21; Gaps 4;

QY 3 VSQPEIRTELGSSAFPCSFNASQGRLAIGSVTWRDEVPKGVRNGTPEFRGLAPL 62

DB 3 VAQPAVLASSRGIAISFVCEY-ASPGKATEVTVTLROADSQVTEVCATMMGNELT-- 59

QY 63 ASSRFLHD-----HQAELHIRDVGHDSIVYCVREVL-----GLGVGTGNGTRLV 108

DB 60 ----FLDSDICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYPYVLGIGNGTQIWI 115

QY 109 VEKEHPQLGAG 119

DB 116 AXEKPSYNRG 126

#### RESULT 10

PCT-US95-15696-2

Sequence 2, Application PC/TUS9515696

GENERAL INFORMATION:

APPLICANT: President and Fellows of Harvard College

TITLE OF INVENTION: SINGLE CHAIN T-CELL RECEPTOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kevin M. Farrell, P.C.

STREET: P.O. Box 999

CITY: York Harbor

STATE: Maine

COUNTRY: USA

ZIP: 03911

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/15696

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/349,915

FILING DATE: 06-DEC-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Farrell, Kevin M.

REGISTRATION NUMBER: 35,505

REFERENCE/DOCKET NUMBER: HU-9404 WO

TELEPHONE: 207-363-0558

TELEFAX: 207-363-0528

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 391 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US95-15696-2

Query Match 14.2%; Score 89.5; DB 5; Length 391;

Best Local Similarity 27.6%; Pred. No. 0.037;

Matches 34; Conservative 17; Mismatches 49; Indels 23; Gaps 7;

QY 6 PPEIRTELGSSAFPCSFNASQGRLAIGSVTWRDEVPKGVRNGTPEFRGR 58

Db 31 POALSIQEGENATKCSYKTS-----INNLOWYRQN--SGRLVHLILIRSNEREKHSR 83  
Qy 59 L-APLASSRFLHDHQAELHIRDVRGHDSIVVCRVEVLGLGVGT-GNGTRLVVEKEHPQL 116  
Db 84 LRVILDTSK-----XSSLLIITATRAADTASIFCATDGGSYIPFGRTSLIV---HPSS 136  
Qy 117 QAG 119  
Db 137 GGG 139

## RESULT 11

US-09-311-784A-32  
; Sequence 32, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermonson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epiimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; FILE REFERENCE: 39563-20022.01  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 32  
; LENGTH: 226  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human MB-1 Ig-alpha  
US-09-311-784A-32

Query Match 13.9%; Score 88; DB 4; Length 226;

Best Local Similarity 29.9%; Pred. No. 0.027;  
Matches 38; Conservative 13; Mismatches 50; Indels 26; Gaps 7;

Qy 1 LWVSO-PPEIRTLGSSAFPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGT--PFRG 57  
Db 33 LWMKVPASLVMVSLGEDAHFQCPENSSNN-----ANVTWR--VLHG----NYTWPPFLG 82  
Qy 58 RLAPLASSRFLHDHQAELHIRDVRGHDSIVVCRVEVLGLGVGTGNGTRLVVEKEHPQ-- 115  
Db 83 P-----GEDPENGTLIIQNVKSHGGIYVCRVQEGNESYQSCGYLRVROPPRPF 133  
Qy 116 --LGAGT 120  
Db 134 LDMGEGT 140

## RESULT 12

PCT-US95-08493-13  
; Sequence 13, Application PC/TUS9508493  
; GENERAL INFORMATION:  
; APPLICANT: Wood, Clive  
; APPLICANT: Caruso, Anthony  
; TITLE OF INVENTION: Novel mlk Receptor Tyrosine Kinases  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LEGAL AFFAIRS  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02140

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/08493  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15234A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-08493-13

Query Match 13.9%; Score 88; DB 5; Length 946;

Best Local Similarity 29.3%; Pred. No. 0.18;  
Matches 27; Conservative 13; Mismatches 24; Indels 28; Gaps 5;

Qy 3 VSQPEIRTLGSSAFPLPCSFNASQGRLAIG-----SVTWFRDEVVPGKEVRNGTPEFRGR 58  
Db 126 IRPDTVRALLGSKVVLPCS-----TMGNPKFAISWFKDETA-----LKNDQP-----R 169  
Qy 59 LAPLASSRFLHDHQAELHIRDVRGHDSIYVC 90  
Db 170 TSVLES-----GNLRIRNVQLEDAGKYRC 193

## RESULT 13

US-09-472-087-102  
; Sequence 102, Application US/09472087  
; Patent No. 6682736  
; GENERAL INFORMATION:  
; APPLICANT: HANSON, DOUGLAS C.  
; APPLICANT: NEVEU, MARK J.  
; APPLICANT: MUELLER, EILLEN E.  
; APPLICANT: HANKE, JEFFREY H.  
; APPLICANT: GILMAN, STEVEN C.  
; APPLICANT: DAVIS, C. GEORFREY  
; APPLICANT: CORVALAN, JOSE R.  
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODIES TO CTLA-4  
; FILE REFERENCE: ABX-PFI  
; CURRENT APPLICATION NUMBER: US/09/472,087  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR FILING DATE: 1999-12-23  
; PRIOR FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 147  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 102  
; LENGTH: 120  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-472-087-102

Query Match 13.8%; Score 87.5; DB 4; Length 120;

Best Local Similarity 27.4%; Pred. No. 0.014;  
Matches 34; Conservative 15; Mismatches 56; Indels 19; Gaps 4;

Qy 3 VSQPEIRTLGSSAFPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGR LAPL 62  
Db 3 VAQPAVVLASSRGIASFVCEY-ASPGKATEVTVLRQADSQVTEVCAATYMGNELT-- 59  
Qy 63 ASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL-----GLGVGTGNGTRLVVE 110

Db 60 ----FLDDSICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYVYLGNGTQIYVI 115  
 QY 111 KEHP 114  
 Db 116 DPEP 119

RESULT 14  
 US-08-067-684-14  
 ; Sequence 14, Application US/08067684  
 ; Patent No. 5434131  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Linsley, Peter S.  
 ; APPLICANT: Ledbetter, Jeffrey A.  
 ; APPLICANT: Damle, Nitin K.  
 ; APPLICANT: Brady, William  
 ; TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sheldon & Mak  
 ; STREET: 225 South Lake Avenue, Suite 900  
 ; CITY: Pasadena  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 91101

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/067,684  
 FILING DATE: 26-May-1993  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Adriano, Sarah B.  
 REGISTRATION NUMBER: 34,470  
 REFERENCE/DOCKET NUMBER: 7848-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 310/312-9900  
 TELEFAX: 310/479-8340  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 187 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-067-684-14

Query Match 13.8%; Score 87.5; DB 1; Length 187;  
 Best Local Similarity 27.4%; Pred. No. 0.024; 56; Indels 19; Gaps 4;  
 Matches 34; Conservative 15; Mismatches 15; Indels 19; Gaps 4;  
 QY 3 VSQPPBIRTEGSSAFPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGR LAPL 62  
 Db 4 VAQPAVTLASSRGIA SFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNELT-- 60  
 QY 63 ASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL---GLGVGTGNGTFLVVE 110  
 Db 61 ----FLDDSICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYVYLGNGTQIYVI 116  
 QY 111 KEHP 114  
 Db 117 DPEP 120

RESULT 15  
 US-08-008-898-14  
 ; Sequence 14, Application US/08008898  
 ; Patent No. 5770197  
 ; GENERAL INFORMATION:

APPLICANT: Linsley, Peter S  
 APPLICANT: Ledbetter, Jeffrey A  
 APPLICANT: Damle, Nitin K  
 APPLICANT: Brady, William  
 TITLE OF INVENTION: CTLA4 RECEPTOR AND METHODS FOR ITS USE  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheldon & Mak  
 STREET: 201 South Lake Avenue, Suite 800  
 CITY: Pasadena  
 STATE: California  
 COUNTRY: United States  
 ZIP: 91101  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/008,898  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/723,617  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Mandel, Saralynn  
 REGISTRATION NUMBER: 31,853  
 REFERENCE/DOCKET NUMBER: 7848  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (818) 796-4000  
 TELEFAX: (818) 795-6321  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 187 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-008-898-14

Query Match 13.8%; Score 87.5; DB 1; Length 187;  
 Best Local Similarity 27.4%; Pred. No. 0.024; 56; Indels 19; Gaps 4;  
 Matches 34; Conservative 15; Mismatches 15; Indels 19; Gaps 4;  
 QY 3 VSQPPBIRTEGSSAFPLPCSFNASQGRLAIGSVTWFRDEVVPGKEVRNGTPEFRGR LAPL 62  
 Db 4 VAQPAVTLASSRGIA SFVCEY-ASPGKATEVRVTVLROADSQVTEVCAATYMMGNELT-- 60  
 QY 63 ASSRFLHD-----HQAELHIRDVRGHDSIYVCRVEVL---GLGVGTGNGTFLVVE 110  
 Db 61 ----FLDDSICTGSSGNQVNLTIQGLRAMDTGLYICKVELMYPYVYLGNGTQIYVI 116  
 QY 111 KEHP 114  
 Db 117 DPEP 120  
 Search completed: February 26, 2004, 12:16:15  
 Job time : 24.2361 secs



Best Local Similarity 62.5%; Pred. No. 10; Mismatches 2; Indels 0; Gaps 0;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19  
||||| :||||| :||  
Db 529 LRAGINKISLSVSVG 544  
||||| :||||| :||

RESULT 3  
T02561  
probable cellulose synthase A2g32620 [imported] - Arabidopsis thaliana  
N;Alternate names: cellulose synthase homolog T26B15.18  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
C;Accession: T02561; D84735  
P;Roundley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,  
submitted to the EMBL Data Library, July 1998  
A;Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.  
A;Reference number: Z14678  
A;Accession: T02561  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-757 <RCU>  
A;Cross-references: EMBL:AC004681; NID:g3298532; PID:g3298550  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Niekman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: D84735  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-757 <STC>  
A;Cross-references: GB:AE002093; NID:g3298550; PIDN:AA25944.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2g32620; T26B15.18  
A;Map position: 2  
A;Introns: 88/2; 194/3; 236/3; 277/3; 344/3; 390/1; 450/3; 566/3

Query Match 52.3%; Score 46; DB 2; Length 757;  
Best Local Similarity 58.8%; Pred. No. 11;  
Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 LRAGFYAVSFLSVAVG 19  
||||| :||||| :||  
Db 740 LKAGFLAVSFVFSVG 756  
||||| :||||| :||

RESULT 4  
T31293  
aldehyde dehydrogenase homolog - Sphingomonas aromaticivorans plasmid pNL1  
C;Species: Sphingomonas aromaticivorans  
C;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C;Accession: T31293  
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G  
submitted to the EMBL Data Library, July 1998  
A;Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A;Reference number: Z20992  
A;Accession: T31293  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-479 <ROW>  
A;Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378434; PIDN:AA004017.1  
C;Genetics:  
A;Genome: plasmid pNL1  
A;Note: nahF  
C;Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match 50.6%; Score 44.5; DB 2; Length 479;  
Best Local Similarity 50.0%; Pred. No. 12;  
Matches 10; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

QY 3 LLRAGF---YAVSFLSVAVG 19  
:||||| :||||| :||  
Db 95 IAKAGFETRAVSLRAAIG 114  
||||| :||||| :||

RESULT 5  
C90157  
hypothetical protein SS00169 [imported] - Sulfolobus solfataricus  
C;Species: Sulfolobus solfataricus  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 24-May-2001  
C;Accession: C90157  
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder,  
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A;Description: Sulfolobus solfataricus complete genome.  
A;Reference number: A99139  
A;Accession: C90157  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-184 <KUR>  
A;Cross-references: GB:AE006641; NID:g13813300; PIDN:AAK40514.1; GSPDB:GN00155  
C;Genetics:  
A;Gene: SS00169

Query Match 50.0%; Score 44; DB 2; Length 184;  
Best Local Similarity 75.0%; Pred. No. 5.7;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 LLRAGFYAVSFL 14  
:||||| :||||| :||  
Db 76 LISAGLYAVSFL 87  
||||| :||||| :||

RESULT 6  
AH1800  
transmembrane efflux protein homolog lin2951 [imported] - Listeria innocua (strain Clif  
C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 06-Jan-2003  
C;Accession: AH1800  
R;Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; M  
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlianc  
A;Title: Comparative Genomics of Listeria species  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AH1800  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-466 <GLA>  
A;Cross-references: GB:AL592022; PIDN:CAC98176.1; PID:g16415492; GSPDB:GN00178  
A;Experimental source: strain Clip11262  
C;Genetics:  
C;Superfamily: multidrug-efflux transporter

Query Match 48.9%; Score 43; DB 2; Length 466;  
Best Local Similarity 58.8%; Pred. No. 21;  
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLRAGFYAVSFLSVAV 18  
||||| :||||| :||  
Db 434 LLTNVGFVSVLSVAV 450  
||||| :||||| :||

RESULT 7  
S03540  
gene frizzled protein precursor - fruit fly (Drosophila melanogaster)  
C;Species: Drosophila melanogaster  
C;Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 02-Mar-2001

C;Species: Mycobacterium leprae  
C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C;Accession: B86918  
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; F  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21126732; PMID:11234002

A;Accession: B86918  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-271 <STO>  
A;Cross-references: GB:AL450380; NID:g13092466; PIDN:CAC29582.1; GSPDB:GN00147  
C;Genetics:  
A;Gene: glpQ

Query Match 47.7%; Score 42; DB 2; Length 271;  
Best Local Similarity 57.9%; Pred.No.18;  
Matches 11; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VLLLRAGFYAVSFLSVAVG 19  
||| ||| : |||  
Db 184 VLLGRAGRYLTSSARTAVG 202

RESULT 10  
B83228  
Probable MPS transporter PA3336 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C;Species: Pseudomonas aeruginosa  
C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C;Accession: B83228  
R;Stover, C.K.; Yoon, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin  
adman, S.; Kuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path  
A;Reference number: AB2950; MUID:20437337; PMID:10984043  
A;Accession: B83228  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-388 <STO>  
A;Cross-references: GB:AEO04756; GB:AEO04091; NID:g9949466; PIDN:AAG06724.1; GSPDB:GN001  
A;Experimental source: strain PAO1  
C;Genetics:  
A;Gene: PA3336  
C;Superfamily: Streptomyces lividans chloramphenicol resistance protein

Query Match 47.7%; Score 42; DB 2; Length 388;  
Best Local Similarity 53.3%; Pred.No.26;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 5 RAGFYAVSFLSVAVG 19  
||| ||| : |||  
Db 157 RASFFAVALVAVLG 171

RESULT 11  
A11426  
transmembrane efflux protein homolog lmo2818 [imported] - Listeria monocytogenes (strai  
C;Species: Listeria monocytogenes  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 06-Jan-2003  
C;Accession: A11426  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A;Authors: Kreft, J.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant  
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant  
A;Title: Comparative genomics of Listeria species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: A11426

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-466 <GLA>  
A;Cross-references: GB:NC 003210; PIDN:CAD01031.1; PID:g16412318; GSPDB:GN00177  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: lmc2818  
C;Superfamily: multidrug-efflux transporter

Query Match 47.7%; Score 42; DB 2; Length 466;  
Best Local Similarity 52.9%; Pred. No. 31;  
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSVAV 18  
DB 434 LTTNAGFCVSLISVAI 450

RESULT 12  
S46096  
Probable membrane protein YBR220c - yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein YBR1510  
C;Species: Saccharomyces cerevisiae  
C;Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 19-Apr-2002  
C;Accession: S46096  
R;Dubois, E.; El Bakloury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.; submitted to the Protein Sequence Database, August 1994  
A;Reference number: S45782  
A;Accession: S46096  
A;Molecule type: DNA  
A;Residues: 1-560 <DUB>  
A;Cross-references: EMBL:D36088; NID:G536609; PID:G536611; GSPDB:GN00002; MIPS:YBR220c  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: MIPS:YBR220c  
A;Cross-references: SGD:S0000424  
A;Map position: 2R  
C;Keywords: transmembrane protein  
F;18-37/Domain: transmembrane #status predicted <TM01>  
F;18-106/Domain: transmembrane #status predicted <TM02>  
F;139-155/Domain: transmembrane #status predicted <TM03>  
F;177-193/Domain: transmembrane #status predicted <TM04>  
F;218-238/Domain: transmembrane #status predicted <TM05>  
F;330-347/Domain: transmembrane #status predicted <TM06>  
F;391-407/Domain: transmembrane #status predicted <TM07>  
F;421-437/Domain: transmembrane #status predicted <TM08>  
F;475-491/Domain: transmembrane #status predicted <TM09>  
F;521-544/Domain: transmembrane #status predicted <TM10>

Query Match 47.7%; Score 42; DB 2; Length 560;  
Best Local Similarity 36.8%; Pred. No. 37;  
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 VLLRAGFYAVSFLSVAV 19  
DB 516 VTLRDGYITNLICIVG 534

RESULT 13  
S74246  
Sulfate transport protein - Arabidopsis thaliana  
N;Alternate names: sulfate transporter  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 28-Oct-1996 #sequence\_revision 14-Nov-1997 #text\_change 20-Jun-2000  
C;Accession: S74246  
R;Takahashi, H.; Sasakura, N.; Noji, M.; Saito, K.  
FEBS Lett. 392, 95-99, 1996  
A;Title: Isolation and characterization of a cDNA encoding a sulfate transporter from Arabidopsis thaliana  
A;Reference number: S74246; MUID:96368029; PMID:8772182  
A;Accession: S74246  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-658 <TAK>

A;Cross-references: EMBL:D85416; NID:g1498119; PID:g1498120  
C;Superfamily: sulfate transport protein  
C;Keywords: sulfate transport; transmembrane protein  
F;92-109/Domain: transmembrane #status predicted <TM1>  
F;119-134/Domain: transmembrane #status predicted <TM2>  
F;140-156/Domain: transmembrane #status predicted <TM3>  
F;172-190/Domain: transmembrane #status predicted <TM4>  
F;194-217/Domain: transmembrane #status predicted <TM5>  
F;253-279/Domain: transmembrane #status predicted <TM6>  
F;278-293/Domain: transmembrane #status predicted <TM7>  
F;338-353/Domain: transmembrane #status predicted <TM8>  
F;371-388/Domain: transmembrane #status predicted <TM9>  
F;408-430/Domain: transmembrane #status predicted <TM10>  
F;435-457/Domain: transmembrane #status predicted <TM11>  
F;464-489/Domain: transmembrane #status predicted <TM12>

Query Match 47.7%; Score 42; DB 2; Length 658;  
Best Local Similarity 50.0%; Pred. No. 43;  
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSVAV 19  
DB 465 LVLIAPFGVLFASVEIG 482

RESULT 14  
F71027  
Hypothetical protein PH1514 - Pyrococcus horikoshii  
C;Species: Pyrococcus horikoshii  
C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
C;Accession: F71027  
R;Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A;Reference number: A71000; MUID:98344137; PMID:9679194  
A;Accession: F71027  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-310 <KAW>  
A;Cross-references: GB:AP000006; NID:G3236133; PIDN:BAA30622.1; PID:G3257939  
A;Experimental source: strain OT3  
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C;Genetics:  
A;Gene: PH1514

Query Match 46.6%; Score 41; DB 2; Length 310;  
Best Local Similarity 58.3%; Pred. No. 30;  
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 8 FYAVSFLSVAV 19  
DB 34 YYALAPLSLGVG 45

RESULT 15  
A70602  
Hypothetical protein Kv0998 - Mycobacterium tuberculosis (strain H37RV)  
C;Species: Mycobacterium tuberculosis  
C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C;Accession: A70602  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A;Reference number: A70500; MUID:98295987; PMID:9634230  
A;Accession: A70602  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-333 <COL>  
A;Cross-references: GB:Z94752; GB:AL123456; NID:g3261731; PIDN:CAB08156.1; PID:el299775

A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: Rv0998

```
Query Match          46.6%; Score 41; DB 2; Length 333;
Best Local Similarity 55.6%; Pred. No. 33;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      2 LLLRAGFYAVSFLSVAVG 19
        :||| | ||||| :|
Db      46 VLLRQGEPAVSFLLISSG 63
```

Search completed: February 26, 2004, 12:14:53  
Job time : 5.87268 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 1.61273 Seconds  
(without alignments)  
613.452 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 88

Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	53.4	835	1 BGAL_LYCES	P48980 lycopersico
2	44	50.0	381	1 O46A DROME	P81919 drosophila
3	43.5	49.4	355	1 S35D HUMAN	Q9ntn3 homo sapien
4	43	48.9	581	1 FRIZ DROME	P18537 drosophila
5	42	47.7	127	1 Y669 YEAST	P38317 saccharomyc
6	42	47.7	217	1 YL02 ARCFU	C28178 archaeoglob
7	42	47.7	412	1 CCA_BUCAP	Q8ka53 buchnera ap
8	42	47.7	560	1 YB70 YEAST	P38318 saccharomyc
9	42	47.7	658	1 Y722 ARATH	P29446 arabidopsis
10	41	46.6	356	1 Y359 AQUAE	O66685 aquifex aeo
11	41	46.6	498	1 YOH2 CAEBEL	Q09502 caenorhabdi
12	40	45.5	123	1 PHF2 RHIME	Q52955 rhizobium m
13	40	45.5	154	1 Y123 ARCFU	C28452 archaeoglob
14	40	45.5	220	1 Y402 PASMU	Q3cmms pasteurilla
15	40	45.5	245	1 DNA_SQUAC	P36983 squalus aca
16	40	45.5	288	1 YD09 SCHPO	O13729 schizosacch
17	40	45.5	434	1 BC13 MOUSE	P59017 mus musculu
18	40	45.5	485	1 BC13 HUMAN	Q9bxx5 homo sapien
19	40	45.5	514	1 YHK8 YEAST	P38776 saccharomyc
20	40	45.5	610	1 YFBS_ECOLI	P77741 escherichia
21	40	45.5	1107	1 ALA2 ARATH	P38205 arabidopsis
22	39	44.3	68	1 ATPJ HUMAN	P56385 homo sapien
23	39	44.3	70	1 ATPJ_PIG	Q9myt8 sus scrofa
24	39	44.3	397	1 GS9C DROME	Q9wlu5 drosophila
25	39	44.3	634	1 S735 ARATH	Q941w6 arabidopsis
26	39	44.3	1221	1 YMP3 CAEBEL	Q10947 caenorhabdi
27	38	43.2	108	1 NTRB_METBA	Q9v8k3 methanosarc
28	38	43.2	199	1 NUGM_MARPO	P26850 cyanidium c
29	38	43.2	201	1 NUGM_CYACA	P48950 cyanidium c
30	38	43.2	207	1 NUGM_PROWI	Q37626 prototheca
31	38	43.2	225	1 TRMA_NEIGO	P55134 reisserichia
32	38	43.2	232	1 CLB3_SHEEP	P80943 ovis aries
33	38	43.2	236	1 BIR8_GORGO	Q95m71 gorilla gor

Q96p09 homo sapien  
Q95m72 pan troglod  
P79826 oncorhynchu  
Q09322 caenorhabdi  
P47434 mycoplasma  
Q9jt82 neisseria m  
P59323 vibrio vuln  
Q07523 bacillus su  
P54428 bacillus su  
P58229 escherichia  
P39183 escherichia  
P32340 saccharomyc

34 38 43.2 236 1 BIR8 HUMAN  
35 38 43.2 236 1 BIR8\_PANTR  
36 38 43.2 258 1 MYPR\_ONCMY  
37 38 43.2 324 1 YR53\_CAEBEL  
38 38 43.2 329 1 Y188\_MYCGB  
39 38 43.2 362 1 TRMA\_NEIMA  
40 38 43.2 403 1 ARGD\_VIBVU  
41 38 43.2 419 1 YHAP\_BACSU  
42 38 43.2 434 1 YRKA\_BACSU  
43 38 43.2 511 1 XASA\_ECO57  
44 38 43.2 511 1 XASA\_ECOLI  
45 38 43.2 513 1 NDII\_YEAST

#### ALIGNMENTS

##### RESULT 1

ID BGAL\_LYCES STANDARD; PRT; 835 AA.  
AC P48980;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Beta-galactosidase precursor (EC 3.2.1.23) (Lactase) (Acid beta-galactosidase) (Exo-(1-->4)-beta-D-galactanase).  
DE Lycopersicon esculentum (tomato)...  
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=cv. Alisa Craig; TISSUE=pericarp;  
RX MEDLINE=95357407; PubMed=7630937;  
RA Carey A.T., Holt K., Picard S., Wilde R., Tucker G.A., Bird C.R.,  
RT "Tomato exo-(1-->4)-beta-D-galactanase. Isolation, changes during  
RT ripening in normal and mutant tomato fruit, and characterization of a  
RT related cDNA clone".  
RL Plant Physiol. 108:1099-1107(1995).  
CC -!- FUNCTION: Involved in cell wall degradation. Degrades  
CC polysaccharides containing beta-(1-->4)-linked galactans, acting  
CC as an exo-(1-->4)-beta-D-galactanase.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-  
CC galactose residues in beta-D-galactosides.  
CC -!- MISCELLANEOUS: Has a pH optimum of 4.5.  
CC -!- SIMILARITY: Belongs to family 35 of glycosyl hydrolases.  
CC -!- SIMILARITY: Contains 1 SUEL-type lectin domain.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X83854; CAA58734.1; -  
PIR; T06590; T06590.  
DR InterPro; IPR008979; Gal\_bind\_like.  
DR InterPro; IPR000922; Gal\_lectin.  
DR InterPro; IPR001944; Glyco\_hydro\_35.  
DR Pfam; PF02140; Gal\_lectin; 1.  
DR Pfam; PF01301; Glyco\_hydro\_35; 1.  
DR PRINTS; PR00742; GLHYDLASE35.  
DR ProDom; PD005612; Gal\_lectin; 1.  
DR PROSITE; PS01182; GLYCOSYL\_HYDROL\_F35; 1.  
DR PROSITE; PS02228; SUEL\_LECTIN; 1.  
DR HydroLase; Glycosidase; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 835 BETA-GALACTOSIDASE.  
FT DOMAIN 749 835 SUEL-TYPE LECTIN.  
FT

FT ACT SITE 180 180 PROTON DONOR (POTENTIAL).  
 FT ACT SITE 249 249 NUCLEOPHILE (POTENTIAL).  
 SQ SEQUENCE 835 AA; 93336 MW; 94C9685F95C4A46 CRC64;  
 Query Match 53.4%; Score 47; DB 1; Length 835;  
 Best Local Similarity 62.5%; Pred. No. 4.7;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 4 LRAGFYAVSFSLVAVG 19  
 DQ 527 LRAGWKISLSTIAGV 542  
 RESULT 2  
 ID\_046A DROME STANDARD; PRT; 381 AA.  
 AC P81919; Q9U5H3.  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Odorant receptor 46a, isoform A.  
 GN OR46A OR OR46B OR OR46F OR OR46F.1 OR OR46F.2 OR DOR46F OR DOR46F.1  
 GN OR DOR19 OR AN8 OR AN9 OR CG17849/CG17849.  
 OS Drosophila melanogaster (Fruit fly)  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=93189723; PubMed=10458908;  
 RA Gao Q.; Chess A.;  
 RT Identification of candidate Drosophila olfactory receptors from  
 RT genomic DNA sequence.;  
 RL Genomics 60:31-39(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G.G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.F., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of Drosophila melanogaster";  
 Science 287:2185-2195(2000).  
 [3]  
 RN REVISIONS.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminler J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review";  
 RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 [4]  
 RN IDENTIFICATION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=99168688; PubMed=10069338;  
 RA Clyne P.J., Warr C.G., Freeman M.R., Lessing D., Kim J., Carlson J.R.;  
 RT "A novel family of divergent seven-transmembrane proteins: candidate  
 odorant receptors in Drosophila";  
 RL Neuron 22:327-338(1999).  
 [5]  
 RN CONCEPTUAL TRANSLATION.  
 RP Robertson H.M.;  
 RA Unpublished observations (MAY-2003).  
 RL -!- FUNCTION: Probable role in the odorant response, being an odorant  
 receptor.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Name=a;  
 CC IsoId=P81919-1; Sequence=Displayed;  
 CC Note=No experimental confirmation available;  
 CC Name=b;  
 CC IsoId=Q9V3N2-1; Sequence=External;  
 CC Note=No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Isoform a is expressed in a subset of 17  
 olfactory receptor neurons in the maxillary palp.  
 CC -!- SIMILARITY: Belongs to family Dr-or of G-protein coupled  
 receptors.  
 CC -!- CAUTION: Ref.2 sequence differs from that shown due to erroneous  
 gene model prediction.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; AE003830; AAF58834.2; ALT\_SEQ.  
 CC FlyBase; F59N002388; Or46a.  
 CC GO; GO:0016021; C:Integral to membrane; NAS.  
 CC GO; GO:0004984; P:Olfactory receptor activity; NAS.  
 CC GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; NAS.  
 CC GO; GO:0007608; P:olfaction; NAS.  
 CC InterPro; IPR004117; 7tm\_6.  
 CC Pfam; PF02949; 7tm\_6; 1.  
 CC Transmembrane; G-protein coupled receptor; Olfaction;  
 CC Alternative splicing; Multigene family.  
 KW DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 38 58 1 (POTENTIAL).  
 FT DOMAIN 59 65 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 66 86 2 (POTENTIAL).  
 FT DOMAIN 87 127 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 128 148 3 (POTENTIAL).  
 FT DOMAIN 149 170 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 171 191 4 (POTENTIAL).  
 FT DOMAIN 192 255 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 256 276 5 (POTENTIAL).  
 FT DOMAIN 277 287 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 288 308 6 (POTENTIAL).  
 FT DOMAIN 309 355 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 356 376 7 (POTENTIAL).  
 FT DOMAIN 377 381 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 381 AA, 44067 MW, 8A2AB426B08F95FA CRC64;  
 Query Match 50.0%; Score 44; DB 1; Length 381;  
 Best Local Similarity 60.0%; Pred. No. 7.3;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 VLLLRAGFYAVSFLS 15  
 ||:|||||:  
 Db 264 VLVLTAFAIAVLIS 278  
 ||:|||||:  
 RESULT 3  
 ID S35D HUMAN STANDARD; PRT; 355 AA.  
 AC Q9NTN3; Q92548;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter (UDP-GlcA/UDP-GalNAc transporter) (Solute carrier family 35 member D1).  
 GN SLC35D1 OR UGTREL7 OR KIAA0260.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC TISSUE=Uterus;  
 RA Muraoka M., Kawakita M., Ishida N.;  
 RX MEDLINE=21223049; PubMed=11322953;  
 RT "Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity."  
 RL FEBS Lett. 495:97-93(2001).  
 [2]  
 SEQUENCE FROM N.A.  
 RC TISSUE=Bone marrow;  
 RX MEDLINE=97191544; PubMed=9039502;  
 RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y., Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. VI. The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by analysis of cDNA clones from cell line KG-1 and brain."  
 RL DNA Res. 3:321-329(1996).  
 [3]  
 SEQUENCE FROM N.A.  
 RA Wilson S.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Transports both UDP-glucuronic acid (UDP-GlcA) and UDP-N-acetylgalactosamine (UDP-GalNAc) from the cytoplasm into the endoplasmic reticulum lumen. May participate in glucuronidation and/or chondroitin sulfate biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic reticulum.  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- SIMILARITY: Belongs to the nucleotide sugar transporter family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; AL133320; CAB92090.1; -  
 DR EMBL; D87449; BAAL1390.1; ALT\_INT.

DR EMBL; AB044343; BAB18586.1; -  
 DR Genew; HGNC:20800; SLC35D1.  
 DR GO; GO:0030176; C:integral to endoplasmic reticulum membrane; NAS.  
 DR GO; GO:0005463; F:UDP-N-acetylgalactosamine transporter activity; NAS.  
 DR GO; GO:0030206; P:chondroitin sulfate biosynthesis; NAS.  
 DR GO; GO:0015789; P:UDP-N-acetylgalactosamine transport; NAS.  
 KW transport; Sugar transport; Transmembrane; Endoplasmic reticulum.  
 FT TRANSMEM 37 59 POTENTIAL.  
 FT TRANSMEM 69 88 POTENTIAL.  
 FT TRANSMEM 158 177 POTENTIAL.  
 FT TRANSMEM 187 205 POTENTIAL.  
 FT TRANSMEM 217 239 POTENTIAL.  
 FT TRANSMEM 254 276 POTENTIAL.  
 FT TRANSMEM 281 303 POTENTIAL.  
 FT TRANSMEM 308 330 POTENTIAL.  
 SQ SEQUENCE 355 AA, 39240 MW, CAFF15D49605CA27 CRC64;  
 Query Match 49.4%; Score 43.5; DB 1; Length 355;  
 Best Local Similarity 70.6%; Pred. No. 8.3;  
 Matches 12; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 QY 3 LLRAGFYAV-SFLSVAV 18  
 |||||:  
 Db 43 LLRAGFYGVSSFLIVV 59  
 |||||:  
 RESULT 4  
 ID FRIZ DROME STANDARD; PRT; 581 AA.  
 AC P19537; Q9VUE0; Q9VUE2;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Frizzled protein precursor (Frizzled-1) (Gfzl).  
 GN FZ OR CGI7697/CG3646.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 [1]  
 SEQUENCE FROM N.A. (ISOFORM I).  
 RX MEDLINE=89159415; PubMed=2493583;  
 RA Vinson C.R., Conover S., Adler P.N.;  
 RT "A Drosophila tissue polarity locus encodes a protein containing seven potential transmembrane domains."  
 RL Nature 338:263-264(1989).  
 [2]  
 SEQUENCE FROM N.A. (ISOFORMS I AND II).  
 RC STRAIN=Canton-S; TISSUE=Embryo;  
 RX MEDLINE=91060073; PubMed=2174014;  
 RA Adler P.N., Vinson C., Park W.J., Conover S., Klein L.;  
 RT "Molecular structure of frizzled, a Drosophila tissue polarity gene."  
 RL Genetics 126:401-416(1990).  
 [3]  
 SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., B. D., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Mays A., Dew I., Dietz S.M., de Pablos B., Delcher A., Deng Z., Mays A., Dugan-Rocha S., Dunn P., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Foeler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
RN [4]  
RP REVISIONS, AND ALTERNATIVE SPLICING.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hradecky P., Huang Y., Kaminler J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
RT systematic review,"  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=Berkley; TISSUE=Embryo;  
RX MEDLINE=22426066; PubMed=12537569;  
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
RA George R.A., Guarini H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,  
RA Rubin G.M., Celinker S.E.;  
RT "A *Drosophila* full-length cDNA resource,"  
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
RN [6]  
RP TOPOLOGY.  
RX MEDLINE=99287588; PubMed=10360843;  
RA Goo J.H., Ahn Y., Park W.J.;  
RT "Frizzled-suc2 fusion gene studies in *Saccharomyces cerevisiae*,"  
RL DNA Cell Biol. 18:429-434(1999).  
CC -1- FUNCTION: Receptor for Wnt proteins. Most of frizzled receptors  
CC are coupled to the beta-catenin canonical signaling pathway, which  
CC leads to the activation of dishevelled proteins, inhibition of  
CC GSK-3 kinase, nuclear accumulation of beta-catenin and activation  
CC of Wnt target genes. A second signaling pathway involving PKC and  
CC calcium fluxes has been seen for some family members, but it is  
CC not yet clear if it represents a distinct pathway or if it can be  
CC integrated in the canonical pathway. As PKC seems to be required  
CC for Wnt-mediated inactivation of GSK-3 kinase. Both pathways seem  
CC to involve interactions of epidermal cells to produce a parallel array  
CC of cuticular hairs and bristles.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=P18537-1; Sequence=Displayed;  
CC Name=1;  
CC -1- DOMAIN: Lys-Thr-X-X-Trp motif is involved in the activation of  
CC the Wnt/beta-catenin signaling pathway (By similarity).  
CC -1- DOMAIN: The FZ domain is involved in binding with Wnt ligands (By  
CC

CC similarity).  
CC -1- SIMILARITY: Belongs to the Fz/Smo G-protein coupled receptor  
CC family.  
CC -1- SIMILARITY: Contains 1 frizzled (FZ) domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X54548; CAA38460.1; -;  
CC EMBL; X54549; CAA38460.1; JOINED.  
CC EMBL; X54550; CAA38460.1; JOINED.  
CC EMBL; X54551; CAA38460.1; JOINED.  
CC EMBL; X54548; CAA38461.1; -;  
CC EMBL; X54549; CAA38461.1; JOINED.  
CC EMBL; X54550; CAA38461.1; JOINED.  
CC EMBL; X54552; CAA38461.1; JOINED.  
CC EMBL; X54546; CAA38458.1; -;  
CC EMBL; X54547; CAA38459.1; ALT\_SEQ.  
CC EMBL; AE003535; AAF49746.3; -;  
CC EMBL; AE003535; AAF49746.3; -;  
CC EMBL; AY051808; AAK93232.1; -;  
CC PIR; S03540; S03540.  
CC FlyBase; FBgn0001085; fz.  
CC GO; GO:0005938; C:cell cortex; IDA.  
CC GO; GO:0016021; C:integral to membrane; TAS.  
CC GO; GO:0005886; C:plasma membrane; IDA.  
CC GO; GO:0037147; F:Wnt-protein binding; ISS.  
CC GO; GO:0045198; P:establishment of epithelial cell polarity; IMP.  
CC GO; GO:0007164; P:establishment of tissue polarity; NAS.  
CC GO; GO:0016318; P:ommatidial rotation; IMP.  
CC GO; GO:0016360; P:sensory organ precursor cell fate determina. .; IMP.  
CC GO; GO:0016055; P:Wnt receptor signaling pathway; ISS.  
CC InterPro; IPR000539; Frizzled.  
CC InterPro; IPR000024; Fz domain.  
CC InterPro; IPR000832; GPCR\_secretin.  
CC Pfam; PF01534; Frizzled; 1.  
CC Pfam; PF01392; Fz; 1.  
CC PRINTS; PR00489; FRIZZLED.  
CC SMART; SM00063; FRI; 1.  
CC PROSITE; PS50038; FZ; 1.  
CC PROSITE; PS50261; G\_PROTEIN\_RECEP\_F2\_4; 1.  
CC Multigene family; Receptor; G-protein coupled receptor; Transmembrane;  
CC Developmental protein; Receptor; Wnt signaling pathway; Glycoprotein; Signal;  
CC Alternative splicing.  
CC SIGNAL 1 18  
CC CHAIN 19 581  
CC DOMAIN 19 247  
CC TRANSMEM 248 270  
CC DOMAIN 271 280  
CC TRANSMEM 281 303  
CC DOMAIN 304 343  
CC TRANSMEM 344 364  
CC DOMAIN 365 380  
CC TRANSMEM 381 401  
CC DOMAIN 402 421  
CC TRANSMEM 422 439  
CC DOMAIN 440 471  
CC TRANSMEM 472 492  
CC DOMAIN 493 529  
CC TRANSMEM 530 553  
CC DOMAIN 554 581  
CC DOMAIN 581 666  
CC SITE 556 561  
CC SITE 579 581  
CC CARBOHYD 67 167  
CC CARBOHYD 167 167  
CC VARSPLIC 406 415  
CC /FTID=vsp\_002013.  
CC  
CC LYS-THR-X-X-X-TRP MOTIF.  
CC PDZ-BINDING.  
CC N-LINKED (GLCNAC. .) (POTENTIAL).  
CC N-LINKED (GLCNAC. .) (POTENTIAL).  
CC DILSGVCFVG -> MYLWQFHTIN (in isoform II).  
CC /FTID=vsp\_002013.



CC accepting and donating subsites within its active site (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: ATP + (tRNA)(N) = diphosphate + (tRNA)(N-1).  
 CC -!- SIMILARITY: Belongs to the tRNA nucleotidyltransferase / poly(A)  
 CC polymerase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE014081; AAM67629.1; -;  
 CC DR InterPro; IPR002646; PolyA\_pol.  
 CC DR Pfam; PF01743; PolyA\_pol; 1.  
 CC KW Transferase; Nucleotidyltransferase; ATP-binding; RNA-binding;  
 CC KW RNA processing; Complete proteome.  
 CC FT NP BIND 68 73 ATP (POTENTIAL).  
 CC SQ SEQUENCE 412 AA; 48592 MW; BOA850ECD5D20C22 CRC64;  
 CC  
 CC Query Match 47.7%; Score 42; DB 1; Length 412;  
 CC Best Local Similarity 43.8%; Pred. No. 17;  
 CC Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
 CC  
 CC QY 3 LLRAGFYAVSFSLVAV 18  
 CC :|||:||||:|  
 CC Db 222 ILFTNFYGISFLSGL 237  
 CC  
 CC RESULT 8  
 CC ID YB70 YEAST STANDARD; PRT; 560 AA.  
 CC AC P38318;  
 CC DT 01-OCT-1994 (Rel. 30, Created)  
 CC DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 CC DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 CC DE Hypothetical 63.0 kDa protein in PYC2-PDB1 intergenic region.  
 CC GN YBR220C OR YBR1510.  
 CC OS Saccharomyces cerevisiae (Baker's yeast).  
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC OX NCBI\_TaxID=4932;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=S288c;  
 CC RA Dubois E., el Bakkoury M., Glansdorff N., Messenguy F., Pierard A.,  
 CC RA Scherens B., Vierendeels F.;  
 CC RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: TO E. COLI AMPG AND TO H. INFLUENZAE HI0350.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z36088; CAA95183.1; -;  
 CC DR FIR; S46096; S46096.  
 CC DR GermOnline; 138763; -;  
 CC DR SGD; S0000424; YBR220C.  
 CC KW Hypothetical protein; Transport; Transmembrane.  
 CC FT TRANSMEM 18 38 POTENTIAL.  
 CC FT TRANSMEM 55 75 POTENTIAL.  
 CC FT TRANSMEM 89 109 POTENTIAL.  
 CC FT TRANSMEM 140 160 POTENTIAL.  
 CC FT TRANSMEM 173 193 POTENTIAL.  
 CC FT TRANSMEM 215 235 POTENTIAL.  
 CC FT TRANSMEM 330 350 POTENTIAL.

FT TRANSMEM 375 395 POTENTIAL.  
 FT TRANSMEM 422 442 POTENTIAL.  
 FT TRANSMEM 522 542 POTENTIAL.  
 SQ SEQUENCE 560 AA; 63009 MW; D10A20FA041B5DF3 CRC64;  
 CC  
 CC Query Match 47.7%; Score 42; DB 1; Length 560;  
 CC Best Local Similarity 36.8%; Pred. No. 22;  
 CC Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
 CC  
 CC QY 1 VLLRAGFYAVSFSLVAVG 19  
 CC :|||:||||:|  
 CC Db 516 VTILRDGYITNLICIVVG 534  
 CC  
 CC RESULT 9  
 CC ID ST22 ARATH STANDARD; PRT; 658 AA.  
 CC AC P92946; O64434; Q8LPG2; Q8VXW4; Q9SH01;  
 CC DT 10-OCT-2003 (Rel. 42, Created)  
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 CC DE Sulfate transporter 2.2 (AST56) (ATH14).  
 CC GN SULTR2;2 OR ATIG7790 OR F28K19.21.  
 CC OS Arabidopsis thaliana (Mouse-ear cress).  
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 CC OX NCBI\_TaxID=3702;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 CC RC STRAIN=cv. Columbia;  
 CC RX MEDLINE=96368029; PubMed=8772182;  
 CC RA Takahashi H., Sasekura N., Noji M., Saito K.;  
 CC RT "Isolation and characterization of a cDNA encoding a sulfate  
 CC RT transporter from Arabidopsis thaliana.";  
 CC RL FEBS Lett. 392:95-99(1996).  
 CC [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=cv. Columbia;  
 CC RX MEDLINE=21016719; PubMed=1130712;  
 CC RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 CC RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
 CC RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
 CC RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
 CC RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 CC RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
 CC RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 CC RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
 CC RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
 CC RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
 CC RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
 CC RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
 CC RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
 CC RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 CC RA Uterback L., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 CC RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 CC RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 CC RT thaliana.";  
 CC RL Nature 408:816-820(2000).  
 CC [3]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=cv. Columbia;  
 CC RX MEDLINE=22954850; PubMed=14593172;  
 CC RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,  
 CC RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,  
 CC RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,  
 CC RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,  
 CC RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,  
 CC RA Arakawa T., Baur J., Banno F., Bowser L., Brooks S.Y., Carninci P.,  
 CC RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjail M., Hansen N.F.,  
 CC RA Hayashizaki Y., Johnson-Hopson C., Huan V.W., Iida K., Karnes M.,  
 CC RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,  
 CC RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,



RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,  
 RT Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.:  
 "Empirical analysis of transcriptional activity in the Arabidopsis  
 genome.";  
 RT Science 302:842-846 (2003).  
 RN [4]

# FUNCTION, TISSUE SPECIFICITY, AND INDUCTION.

RP MEDLINE=20387013; PubMed=10929111;  
 RX Takahashi H., Watanabe-Takahashi A., Smith F.W., Blake-Kalff M.,  
 RA Hawkeford M.J., Saito K.;  
 RA "The roles of three functional sulphate transporters involved in  
 RT uptake and translocation of sulphate in Arabidopsis thaliana.";  
 RL Plant J. 23:171-182 (2000).  
 RN [5]

# INDUCTION.

RP MEDLINE=22402754; PubMed=12514246;  
 RX Okama N., Takei K., Sakakibara H., Hayashi H., Yoneyama T.,  
 RA Fujiwara T.;  
 RA "Regulation of sulfur-responsive gene expression by exogenously  
 RT applied cytokinins in Arabidopsis thaliana.";  
 RL Plant Cell Physiol. 43:1493-1501 (2002).

CC -1- FUNCTION: Low-affinity H(+)/sulfate cotransporter that may be  
 involved in the distribution of sulfate from vascular bundles to  
 the palisade cells of the leaves. Plays a central role in the  
 regulation of sulfate assimilation.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- TISSUE SPECIFICITY: Expressed in the phloem in roots and in  
 the phloem of vascular bundles in leaves.  
 CC -1- INDUCTION: In leaves by sulfate starvation. Up-regulated after  
 treatment with zeatin, an exogenous cytokinin.  
 CC -1- SIMILARITY: Belongs to the SLC26A/Sulp transporter (TC 2.A.53)  
 family.

CC -1- SIMILARITY: Contains 1 STAS domain.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous  
 gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D85416; BAA12811.1; --  
 DR EMBL; AB012047; BAA25174.1; --  
 DR EMBL; AC009243; AAF17693.1; ALT\_SEQ.  
 DR EMBL; AY099863; AAM20714.1; --  
 DR EMBL; AY074516; AAL67130.2; --  
 DR PIR; S74246; S74246.  
 DR InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulph\_transpt.  
 DR Pfam; PF01740; STAS; 1.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRfam; TIGR00815; sulp; 1.  
 DR PROSITE; PS01130; SLC26A; 1.  
 DR PROSITE; PS50801; STAS; 1.  
 KW Transport; Symport; Sulfate transport; Transmembrane; Glycoprotein;  
 KW Multigene family.  
 FT DOMAIN 1 91 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 92 112 POTENTIAL.  
 FT DOMAIN 113 114 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 115 135 POTENTIAL.  
 FT DOMAIN 136 139 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 140 160 POTENTIAL.  
 FT DOMAIN 161 171 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 172 192 POTENTIAL.  
 FT DOMAIN 193 194 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 195 215 POTENTIAL.  
 FT DOMAIN 216 251 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 252 272 POTENTIAL.  
 FT DOMAIN 273 277 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 278 298 POTENTIAL.

FT DOMAIN 299 333 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 334 354 POTENTIAL.  
 FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 371 391 POTENTIAL.  
 FT DOMAIN 392 403 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 404 424 POTENTIAL.  
 FT DOMAIN 425 427 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 428 448 POTENTIAL.  
 FT DOMAIN 449 463 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 464 484 POTENTIAL.  
 FT DOMAIN 485 658 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 659 687 STAS.  
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 89 89 K -> R (IN REF. 1; BAA12811).  
 FT CONFLICT 115 115 L -> P (IN REF. 3; AAL67130).  
 FT CONFLICT 584 584 V -> M (IN REF. 1; BAA25174).  
 SQ SEQUENCE 658 AA; 72317 MW; 587A1D0FD5462FCF CRC64;

Query Match 47.7%; Score 42; DB 1; Length 658;  
 Best Local Similarity 50.0%; Pred. No. 25;  
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSVAVG 19

DB 465 LVLIARFGVLFASVEIG 482

# RESULT 10

# Y359 AQUAE

ID Y359 AQUAE STANDARD; PRT; 356 AA.  
 AC O66685;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein AQ\_359.

# GN AQ\_359.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.

OX NCBI\_TaxID=63363;

RN [1]\_TaxID=63363;

RP SEQUENCE FROM N.A.

RC STRAIN=VFS;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus.";

RL Nature 392:353-358 (1998).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB000685; AAC06646.1; --

DR PIR; A70332; A70332.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 7 29 POTENTIAL.

FT TRANSMEM 49 71 POTENTIAL.

FT TRANSMEM 91 113 POTENTIAL.

FT TRANSMEM 270 292 POTENTIAL.

FT TRANSMEM 299 316 POTENTIAL.

FT TRANSMEM 329 348 POTENTIAL.

SQ SEQUENCE 356 AA; 41807 MW; 4E9EA78C38132D05 CRC64;

# Query Match

46.6%; Score 41; DB 1; Length 356;

Best Local Similarity 37.5%; Pred. No. 21;

Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 LLIRAGFYAVSLSVAV 18  
ID YQH2 CAEEL STANDARD; PRT; 498 AA.  
DB 53 ILPLGFFYISFILTSLI 68

RESULT 11

YQH2 CAEEL  
ID YQH2 CAEEL STANDARD; PRT; 498 AA.  
AC Q09502; (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Probable G protein-coupled receptor C43C3.2.  
GN C43C3.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN Sultston J.;  
RP SEQUENCE FROM N.A.  
RA STRAIN=Bristol N2;  
RC Sultston J.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Not known. Putative receptor.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
CC MOST SIMILAR TO SOMATOSTATIN RECEPTORS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z47067; CAA87329.1; -  
CC PIR; T19901; T19901.  
CC WormPep; C43C3.2; C501524.  
CC InterPro; IPR000276; GPCR\_Rhodopsn.  
CC Pfam; PF00001; 7tm1; 1.  
CC PRINTS; PR00237; GPCR\_Rhodopsn.  
CC PROSITE; PS00237; G-PROTEIN RECF\_F1\_1; 1.  
CC PROSITE; PS00262; G-PROTEIN RECF\_F1\_2; 1.  
CC Hypothetical protein; G-protein coupled receptor; Transmembrane;  
KW Glycoprotein.  
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 53 73 1 (POTENTIAL).  
FT DOMAIN 74 93 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 94 114 2 (POTENTIAL).  
FT DOMAIN 115 130 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 131 151 3 (POTENTIAL).  
FT DOMAIN 152 170 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 171 191 4 (POTENTIAL).  
FT DOMAIN 192 258 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 259 279 5 (POTENTIAL).  
FT DOMAIN 280 313 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 314 334 6 (POTENTIAL).  
FT DOMAIN 335 356 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 357 377 7 (POTENTIAL).  
FT DOMAIN 378 498 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 446 450 POLY-ASN.  
FT CARBOHYD 48 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 498 AA; 57463 MW; 7F13584348FB546 CRC64;

Query Match 46.6%; Score 41; DB 1; Length 498;

Best Local Similarity 61.5%; Pred. No. 29; Mismatches 0; Gaps 0;  
Matches 8; Conservative 3; Indels 2

QY 2 LLIRAGFYAVSLSVAV 14  
ID YQH2 CAEEL STANDARD; PRT; 498 AA.  
DB 53 ILPLGFFYISFILTSLI 68

Db 51 LMLIAGFYALLFM 63

RESULT 12

PHF2\_RHIME  
ID PHF2\_RHIME STANDARD; PRT; 123 AA.  
AC Q52965;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Phaf2 protein.  
GN PHAF2 OR R00397 OR SMC00056.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=RCR2011 / SU47;  
RX MEDLINE=95276304; PubMed=7756693;  
RA Kellier M., Roxlau A., Weng W.M., Schmidt M., Quandt J., Niehaus K.,  
RA Jording D., Arnold W., Puchler A.;  
RT "Molecular analysis of the Rhizobium meliloti mucR gene regulating the  
RT biosynthesis of the exopolysaccharides succinoglycan and  
RT galactoglucan";  
RL Mol. Plant Microbe Interact. 8:267-277 (1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA STRAIN=1021;  
RX MEDLINE=21396507; PubMed=11481430;  
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Giotard T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
RA Pohl T., Portetelle D., Puchler A., Purnelle B., Rampsperger U.,  
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
RT "Analysis of the chromosome sequence of the legume symbiont  
RT Sinorhizobium meliloti strain 1021";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: BELONGS TO THE PHAF/MRPF FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L37353; AAA74237.1; -  
CC DR EMBL; AL591785; CAC45569.1; -  
CC DR InterPro; IPR007208; Mrpf\_Phaf.  
CC Pfam; PF04066; Mrpf\_Phaf; 1.  
CC Transport; Potassium transport; Transmembrane; Complete proteome.  
KW TRANSMEM 4 24 POTENTIAL.  
FT TRANSMEM 38 58 POTENTIAL.  
FT TRANSMEM 64 84 POTENTIAL.  
SQ SEQUENCE 123 AA; 12968 MW; 86672D8DB43D0A6E CRC64;  
Query Match 45.5%; Score 40; DB 1; Length 123;  
Best Local Similarity 36.8%; Pred. No. 12;  
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;  
QY 1 VLLIRAGFYAVSLSVAV 19  
ID YQH2 CAEEL STANDARD; PRT; 498 AA.  
DB 53 ILPLGFFYISFILTSLI 68

RESULT 13

YI23\_ARCFU  
ID YI23\_ARCFU STANDARD; PRT; 154 AA.  
AC O28452;  
DT 28-FEB-2003 (Rel. 41, Created)



```

DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein AF1823.
GN AF1823.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OC NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=839475;
RA Klenk H.-P., Clayton R.A., Tomb J.-P., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Karpavich A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AE000976; AAB89426.1; -.
DR F1R; F69477; F69477.
DR TIGR; AF1823; -.
DR HYPOTHETICAL protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 24 POTENTIAL.
FT TRANSMEM 29 48 POTENTIAL.
FT TRANSMEM 53 75 POTENTIAL.
FT TRANSMEM 87 109 POTENTIAL.
FT TRANSMEM 124 146 POTENTIAL.
FT SEQUENCE 154 AA; 16504 MW; 2602FA2DDBD060EA CRC64;
SQ
Query Match 45.5%; Score 40; DB 1; Length 154;
Best Local Similarity 44.4%; Pred. No. 15;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSLVAV 18
DB 22 VLLTKDNFYALYNSVTM 39
|||||:|||||:|||||:
|||||:|||||:|||||:

RESULT 14
Y402 PASWU
ID Y402 PASWU STANDARD; PRT; 220 AA.
AC Q9CWM5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PM0402.
GN PM0402.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OC NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RA "Complete genomic sequence of Pasteurella multocida PM70."
RA Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the BII family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AE006076; AAK02486.1; -.
DR InterPro; IPR006214; UPF0005.
DR Pfam; PF01027; UPF0005.1
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 74 94 POTENTIAL.
FT TRANSMEM 105 125 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 196 216 POTENTIAL.
FT SEQUENCE 220 AA; 23936 MW; 8EECE2E3FB4B5CDD5 CRC64;
SQ
Query Match 45.5%; Score 40; DB 1; Length 220;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSL 15
DB 52 LILMLAGFYGLLFLT 66
|||||:|||||:|||||:
|||||:|||||:|||||:

RESULT 15
DMA SQAC
ID DMA SQAC STANDARD; PRT; 245 AA.
AC F16963;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Proteolipid protein DM alpha.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squalidae; Squalus.
OC NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94000810; PubMed=8398138;
RA Kitagawa K., Sinoway M.P., Yang C., Gould R.M., Colman D.R.;
RA "A proteolipid protein gene family: expression in sharks and rays and
RA possible evolution from an ancestral gene encoding a pore-forming
RA polypeptide."
RT Neuron 11:433-448(1993).
RL Neuron 11:433-448(1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN WHITE MATTER IN
CC MYELINATING SHARK BRAIN.
CC -1- SIMILARITY: Belongs to the myelin proteolipid protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U02973; AAC59639.1; -.
DR F1R; I51323; I51323.
DR InterPro; IPR001614; Myelin_PLP.

```

DR Pfam, PF01275; Myelin\_PLP; 1.  
DR PRINTS; PRO0214; MYELINPLP.  
DR SMART; SM00002; PLP; 1.  
DR PROSITE; PS00575; MYELIN\_PLP\_1; 1.  
DR PROSITE; PS01004; MYELIN\_PLP\_2; 1.  
KW Myelin; Transmembrane.  
FT TRANSMEM 19 35 POTENTIAL.  
FT TRANSMEM 71 87 POTENTIAL.  
FT TRANSMEM 117 133 POTENTIAL.  
FT TRANSMEM 204 220 POTENTIAL.  
SQ SEQUENCE 245 AA; 26910 MW; 521B0A11076B16C5 CRC64;  
  
Query Match 45.5%; Score 40; DB 1; Length 245;  
Best Local Similarity 47.4%; Pred. No. 22;  
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;  
  
Cy 1 VLLIRAGFYAVSFISVAVG 19  
||| | | : |  
Db 84 VLLIAGFYTTSAVKALFG 102

Search completed: February 26, 2004, 12:10:49  
Job time : 3.61273 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 8.11406 Seconds  
(without alignments)  
738.822 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 88

Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	151	6 Q8MJ01	Q8MJ01 macaca mula
2	88	100.0	152	4 Q95668	Q95668 homo sapien
3	88	100.0	165	4 Q95669	Q95669 homo sapien
4	88	100.0	176	4 Q95667	Q95667 homo sapien
5	88	100.0	176	6 Q95J88	Q95J88 macaca fasc
6	88	100.0	177	4 Q14930	Q14930 homo sapien
7	88	100.0	190	4 Q14932	Q14932 homo sapien
8	88	100.0	201	4 Q14931	Q14931 homo sapien
9	85	96.6	180	6 Q8MJ02	Q8MJ02 macaca mula
10	70	79.5	192	11 Q8CG11	Q8CG11 rattus norv
11	70	79.5	192	11 Q8OWM8	Q8OWM8 rattus norv
12	70	79.5	192	11 Q8CFD9	Q8CFD9 rattus norv
13	51	58.0	847	10 Q9SCW1	Q9SCW1 arabidopsis
14	51	58.0	847	10 Q8RWC1	Q8RWC1 arabidopsis
15	49	55.7	610	16 Q7VEF7	Q7VEF7 prochloroco
16	48	54.5	707	10 Q65761	Q65761 cicer ariet

17	48	54.5	721	10 Q9ZP30	Q9ZP30 carica papa
18	47	53.4	838	10 Q9ZP11	Q9ZP11 lycopersico
19	46	52.3	724	10 Q9TOP6	Q9TOP6 lycopersico
20	46	52.3	724	10 Q81100	Q81100 lycopersico
21	46	52.3	757	10 Q80899	Q80899 arabidopsis
22	46	52.3	3763	5 Q8T2A1	Q8T2A1 dictyosteli
23	45	51.1	230	10 Q7X9C6	Q7X9C6 pyrus pyrif
24	45	51.1	270	16 Q8RD77	Q8RD77 thermomane
25	45	51.1	376	10 Q9FVH5	Q9FVH5 prunus arme
26	45	51.1	843	10 Q93X58	Q93X58 fragaria an
27	45	51.1	1561	16 Q888F2	Q888F2 pseudomonas
28	44.5	50.6	479	2 Q86001	Q86001 sphingomona
29	44	50.0	184	17 Q980V8	Q980V8 sulfolobus
30	44	50.0	351	16 Q8G3L8	Q8G3L8 bifidobacte
31	44	50.0	476	16 Q82WR1	Q82WR1 nitrosomona
32	44	50.0	591	16 Q8P300	Q8P300 streptococc
33	44	50.0	591	16 Q8K8X7	Q8K8X7 streptococc
34	44	50.0	602	16 Q8CR29	Q8CR29 staphylococ
35	44	50.0	739	10 Q9M5J3	Q9M5J3 phaseolus a
36	44	50.0	845	10 Q9LLS9	Q9LLS9 lycopersico
37	43.5	49.4	202	11 Q8BWB7	Q8BWB7 mus musculu
38	43.5	49.4	306	11 Q8BX24	Q8BX24 mus musculu
39	43.5	49.4	334	11 Q8BXX2	Q8BXX2 mus musculu
40	43	48.9	160	8 Q8M0F0	Q8M0F0 cryptococcu
41	43	48.9	202	8 Q8M0E9	Q8M0E9 cryptococcu
42	43	48.9	202	8 Q85T00	Q85T00 cryptococcu
43	43	48.9	294	17 Q8TQJ7	Q8TQJ7 methanosarc
44	43	48.9	338	17 Q96Y76	Q96Y76 sulfolobus
45	43	48.9	466	16 Q926T9	Q926T9 listeria in

## ALIGNMENTS

### RESULT 1

Q8MJ01	PRELIMINARY;	PRT;	151 AA.
AC Q8MJ01;			
DT 01-OCT-2002 (TRENBLrel. 22, Created)			
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)			
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)			
DE NKP30VIED.			
OS Macaca mulatta (Rhesus macaque).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;			
OC Cercopitheidae; Macaca.			
OX NCBI_TaxID=9544;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA LaBonte M.L., Miller J., Letvin N.L.;			
RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification			
RT of NKP46SD and NKP30S."			
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY035216; AAK63118.1; --			
DR InterPro; IPR003599; IG-			
DR InterPro; IPR007110; IG-like.			
DR Pfam; PF00047; IG; 1.			
DR SMART; SM00409; IG; 1.			
DR PROSITE; PS00835; IG LIKE; 1.			
SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;			

Query Match 100.0%; Score 88; DB 6; Length 151;

Best Local Similarity 100.0%; Pred. No. 1.8e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFLSVAVG 19

DB 114 VLLLRAGFYAVSFLSVAVG 132

### RESULT 2

Q95668

ID Q95668 PRELIMINARY; PRT; 152 AA.

AC O95668;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 1C7f.  
 GN GN  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tone P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NFkappaB  
 RT family within a 90 kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=86293202;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=75909564;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierrina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., Demarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
 RT Level of TNF-beta Production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53(1991).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";

RL Hum. Mol. Genet. 3:793-799(1994).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vainan M.,  
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
 RT Nuclear RNA Helicase of the D-E-A-D Family.";  
 RL Genomics 26:210-218(1995).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing.";  
 RL Genomics 45:591-600(1997).  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT LkBL and MICA genes at the centromeric end of the HLA Class I  
 RT region.";  
 RL Genomics 47:372-382(1998).  
 DR EMBL; Y14768; CAA75067.1; .  
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG LIKE; 1.  
 DR SEQUENCE 152 AA; 16393 MW; 42718746451P9ADC CRC64;  
 SQ  
 Query Match 100.0%; Score 88; DB 4; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VLLLRAGFYAVSFLSVAVG 19  
 DB 114 VLLLRAGFYAVSFLSVAVG 132  
 ID O95669 PRELIMINARY; PRT; 165 AA.  
 AC O95669;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 1C7f.  
 GN GN  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,

RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NFKappaB  
 RL family within a 90 kilobase HLA class III segment."; Nat. Genet. 3:137-145(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcenci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 RL molecule expressed in transplanted human hearts."; Transplantation 61:1387-1392(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RL TNF region."; Immunogenetics 42:315-322(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a  
 RL heteromeric complex with lymphotoxin on the cell surface."; Cell 72:847-856(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pemica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RL homology and chromosomal localization."; Nucleic Acids Res. 13:6361-6373(1985).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
 RL Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
 RL Level of TNF-beta Production."; J. Exp. Med. 173:209-219(1991).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene."; Immunogenetics 33:50-53(1991).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RL histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins."; Hum. Mol. Genet. 3:793-799(1994).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
 RL Nuclear RNA Helicase of the D-E-A-D Family."; Genomics 26:210-218(1995).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;

RT "Alternative splicing of the LST-1 gene located in the major  
 RL histocompatibility complex on human chromosome 6."; DNA Seq. 8:155-160(1997).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Mater S., Martinozzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RL differential regulation, initiation, and alternative splicing."; Genomics 45:591-600(1997).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RL LkBL and MICA genes at the centromeric end of the HLA Class I  
 RL region."; Genomics 47:372-382(1998).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=914768; PubMed=750681;  
 RA GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG-like.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS00835; IG LIKE; 1.  
 SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;  
 Query Match 100.0%; Score 88; DB 4; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 2e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VLLLRAGFYAVSFLSVAVG 19  
 |||||  
 DB 114 VLLLRAGFYAVSFLSVAVG 132  
 RESULT 4  
 O95667 PRELIMINARY; PRT; 176 AA.  
 ID O95667  
 AC O95667; DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 1C7e.  
 GN 1C7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NFKappaB  
 RL family within a 90 kilobase HLA class III segment."; Nat. Genet. 3:137-145(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcenci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 RL molecule expressed in transplanted human hearts."; Transplantation 61:1387-1392(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=96006555; PubMed=7590964;  
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
RA Weiss E.H.;  
RT "Cloning and genomic characterization of LST1: a new gene in the human  
RT TNF region.";  
RL Immunogenetics 42:315-322(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93208881; PubMed=7916555;  
RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
RT "Lymphotoxin-beta: A new member of the TNF family that forms a  
RT heteromeric complex with lymphotoxin on the cell surface.";  
RL Cell 72:847-856(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96016093; PubMed=2995927;  
RA Nedwin G.E., Naylor S.D., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
RA Pennica D., Goeddel D.V., Gray P.W.;  
RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
RT homology and chromosomal localization.";  
RL Nucleic Acids Res. 13:6361-6373(1985).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91086846; PubMed=1670638;  
RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
RA Riethmüller G., Weiss E.H.;  
RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
RT Level of TNF-beta Production.";  
RL J. Exp. Med. 173:209-219(1991).  
RN [7]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91139175; PubMed=1671667;  
RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
RT "Haplotypic polymorphisms of the TNFB gene.";  
RL Immunogenetics 33:50-53(1991).  
RN [8]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94362679; PubMed=8081366;  
RA Albertella M.R., Campbell D.R.;  
RT "Characterization of a novel gene in the human major  
RT histocompatibility complex that encodes a potential new member of the  
RT I kappa B family of proteins.";  
RL Hum. Mol. Genet. 3:793-799(1994).  
RN [9]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95324911; PubMed=7601445;  
RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
RA Strominger J., Spies T.;  
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
RT Nuclear RNA Helicase of the D-E-A-D Family.";  
RL Genomics 26:210-218(1995).  
RN [10]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20132445; PubMed=10668961;  
RA Neville M.J., Campbell R.D.;  
RT "Alternative splicing of the LST-1 gene located in the major  
RT histocompatibility complex on human chromosome 6.";  
RL DNA Seq. 8:153-160(1997).  
RN [11]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98035883; PubMed=9367684;  
RA de Baey A., Fellerhoff B., Mater S., Martinozzi S., Weidle U.,  
RA Weiss E.H.;  
RT "Complex expression pattern of the TNF region gene LST1 through  
RT differential regulation, initiation, and alternative splicing.";  
RL Genomics 45:591-600(1997).  
RN [12]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98149985; PubMed=9480751;

RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
RA Kimura M., Inoko H.;  
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
RT IKBL and MICA genes at the centromeric end of the HLA class I  
RT region.";  
RL Genomics 47:372-382(1998).  
DR EMBL: Y14768; CAA75066.1; --  
DR GO: GO:0003793; P:defense/immunity protein activity; NAS.  
DR GO: GO:0006955; P:immune response; NAS.  
DR GO: GO:0006954; P:inflammatory response; NAS.  
DR InterPro: IPR003599; IG.  
DR InterPro: IPR007110; IG-like.  
DR Pfam: PF00047; IG; 1.  
DR SMART: SM00409; IG; 1.  
DR PROSITE: PS0835; IG LIKE; 1.  
SQ SEQUENCE 176 AA; 18749 MW; 162BB775DA2BCD35 CRC64;  
Query Match 100.0%; Score 88; DB 4; Length 176;  
Best Local Similarity 100.0%; Pred. No. 2.le-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLLLRAGFYAVSFSLVAVG 19  
DB 114 VLLLRAGFYAVSFSLVAVG 132  
Q95JB8 PRELIMINARY; PRT; 176 AA.  
AC Q95JB8; --  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE NKP30 (NKP30v1).  
GN NCR.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and  
OC Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9541, 9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.fascicularis; TISSUE=Lymphoid;  
RA Rizzi M., Biasoni R.;  
RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in  
RT Macaca fascicularis lymphoid cells.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.mulatta;  
RA LaBonte M.L., Miller J., Letvin N.L.;  
RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification  
RT of NKP46SD and NKP30S.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ278389; CAC41081.1; --  
DR EMBL: AY035215; AK63117.1; --  
DR InterPro: IPR007110; IG-like.  
DR Pfam: PF00047; IG; 1.  
DR PROSITE: PS0835; IG LIKE; 1.  
SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;  
Query Match 100.0%; Score 88; DB 6; Length 176;  
Best Local Similarity 100.0%; Pred. No. 2.le-06;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLLLRAGFYAVSFSLVAVG 19  
DB 139 VLLLRAGFYAVSFSLVAVG 157



FT SIGNAL 1 61 POTENTIAL.  
 FT CHAIN 62 177 1C7.  
 SQ SEQUENCE 177 AA; 19237 MW; DD5EC96F0AB2DCE6 CRC64;  
 Query Match 100.0%; Score 88; DB 4; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLLLRAGFYAVSFSLVAVG 19  
 |||||  
 Db 139 VLLLRAGFYAVSFSLVAVG 157

RESULT 7  
 OL14932 PRELIMINARY; PRT; 190 AA.  
 ID OL14932 AC OL14932;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE 1C7 precursor (Natural killer cell receptor).  
 GN 1C7 OR NKFP30.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=98422187; PubMed=8824804;  
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;  
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."  
 RL Genomics 31:215-222(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX Nalabolu S.R., Raghunathan A., Weissman S.M.;  
 RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bouqueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tone P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NF kappa B  
 RT family within a 90 Kilobase HLA class III segment."  
 RL Nat. Genet. 3:137-145(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcoci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts."  
 RL Transplantation 61:1387-1392(1996).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region."  
 RL Immunogenetics 42:315-322(1995).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface."  
 RL Cell 72:847-856(1993).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;

RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization."  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production."  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene."  
 RL Immunogenetics 33:50-53(1991).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins."  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nures M., Renard C., Geffrotin C., Vainan M.,  
 RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family."  
 RL Genomics 26:210-218(1995).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6."  
 RL DNA Seq. 8:155-160(1997).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing."  
 RL Genomics 45:591-600(1997).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IkBL and MICA genes at the centromeric end of the HLA class I  
 RT region."  
 RL Genomics 47:372-382(1998).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymphoid;  
 RA Biassoni R., Pessino A., Malaspina A.;  
 RT "NK-A1 activating NK receptor."  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Peripheral blood;



RA Sato M., Yabe T., Ohashi J., Tsuchiya N., Hanaoka K., Tokunaga K.,  
 RA Juji T.;  
 RA "Identification of two novel single nucleotide polymorphisms in the  
 RT Nkp30 gene in human natural killer cells.";   
 RT Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF031138; AA866590.1; -;  
 DR EMBL; Y14768; CAA75065.1; -;  
 DR EMBL; AJ223153; CAB54004.1; -;  
 DR EMBL; AB055881; BAB78472.1; -;  
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SMO0409; IG; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 KW Signal; Receptor.  
 FT SIGNAL 1 61 POTENTIAL.  
 FT CHAIN 62 190 IC7.  
 SQ SEQUENCE 190 AA; 20640 MW; 1FAC919E20A6B19A CRC64;  
 Query Match 100.0%; Score 88; DB 4; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 VLLLRAGFAYVFLSVAVG 19  
 |||||  
 Db 139 VLLLRAGFAYVFLSVAVG 157  
 |||||  
 RESULT 8  
 OL4931  
 ID OL4931 PRELIMINARY; PRT; 201 AA.  
 AC OL4931;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE IC7 precursor (NCR3 protein).  
 GN IC7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=96422187; PubMed=8824804;  
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;  
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."  
 RL Genomics 31:215-222(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RA Nalabolu S.R., Raghunathan A., Weissman S.M.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,  
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,  
 RA Laaky S., Hood L.;  
 RT "Sequence of the human major histocompatibility complex class III  
 RT region.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=93272029; PubMed=84999847;  
 RA Iris F., Souguieret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NF kappa B  
 RT family within a 90 kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN [5]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcesi R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., Demarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2395927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Page G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53(1991).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geoffrotin C., Vaiman M.,  
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Stroninger J., Spies T.;  
 RT "The BAV1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family.";  
 RL Genomics 26:210-218(1995).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN [14]  
 RP SEQUENCE FROM N.A.

```

Q8MJ02
ID Q8MJ02 PRELIMINARY; PRT; 180 AA.
AC Q8MJ02;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NKP30.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OC NCBI_TaxID=9544;
OX [1]
RN RNP SEQUENCE FROM N.A.
RA LaBonte M.L.; Miller J.; Letvin N.L.;
RT "Molecular cloning of rhesus monkey NKP46 and NKP30 and identification
of NKP46SD and NKP30S.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY035214; AAK63116.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;
SQ
Query Match 96.6%; Score 85; DB 6; Length 180;
Best Local Similarity 94.7%; Pred. No. 6.7e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLLLRAGFYAVSFLSVAG 19
Db 139 VLLLRAGFYAVSFLSVAG 157
RESULT 10
Q8CG11
ID Q8CG11 PRELIMINARY; PRT; 192 AA.
AC Q8CG11;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NK receptor 1c7 precursor.
GN 1c7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
OX [1]
RN RNP SEQUENCE FROM N.A.
RA STRAIN=BN.
RA Backman-Petersson E.; Miller J.R.; Hollyoake M.; Aguado E.;
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor 1c7.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430419; CAD23067.2; -.
DR EMBL; AJ430420; CAD23067.2; JOINED.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.

```

```

KW Signal; Receptor.      19 POTENTIAL.
FT SIGNAL                1
SQ SEQUENCE 192 AA; 20470 MW; 439AD7AGAFBE6DC0 CRC64;

Query Match      79.5%; Score 70; DB 11; Length 192;
Best Local Similarity 83.3%; Pred. No. 0.002;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFLSVAVG 19
||||| |||:|||||

```

```

Db      148 LLLRAGVYALSFSLSVATG 165

RESULT 11
Q80WM8 PRELIMINARY; PRT; 192 AA.
AC Q80WM8;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NRP30.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW;
RX MEDLINE=2216811; PubMed=12180816;
RA Hsieh C.L., Obara H., Ogura Y., Martinez O.M., Krams S.M.;
RT "NK cells and transplantation.";
RL Transpl. Immunol. 9:111-114(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW;
RA Haieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RT "Identification, Cloning, and Characterization of a Novel Rat NK
RT Receptor, rNKP30: a Molecule Expressed in Liver Allografts.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY273824; AAP13457.1; -
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 192 AA; 20498 MW; 7DF58B245C52377 CRC64;

Query Match 79.5%; Score 70; DB 11; Length 192;
Best Local Similarity 83.3%; Pred. No. 0.002;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFSLSVAVG 19
|||||:|:|:|:|
Db      148 LLLRAGVYALSFSLSVATG 165

RESULT 12
Q8CFD9 PRELIMINARY; PRT; 192 AA.
AC Q8CFD9;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE 1c7 protein precursor.
GN 1c7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=NK cell;
RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RT "Molecular characterization of the novel rat NK receptor 1c7.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430418; CAD23066.1; -
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.

Query Match 79.5%; Score 70; DB 11; Length 192;
Best Local Similarity 83.3%; Pred. No. 0.002;
Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 LLLRAGFYAVSFSLSVAVG 19
|||||:|:|:|:|
Db      148 LLLRAGVYALSFSLSVATG 165

RESULT 13
Q9SCW1 PRELIMINARY; PRT; 847 AA.
AC Q9SCW1;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative beta-galactosidase precursor (EC 3.2.1.23)
DE (Lactase).
GN BGAL1
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Gy I., Kreis M., Lecharny A.;
RT "the Beta-galactosidases are encoding by a multigene family in
RT Arabidopsis thaliana.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
CC GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
DR EMBL; AJ270297; CAB64737.1; -
DR EMBL; AP001307; BAB01923.1; -
DR GO; GO:0009341; C:beta-galactosidase complex; IEA.
DR GO; GO:0004565; F:beta-galactosidase activity; IEA.
DR GO; GO:0016738; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0005975; F:carbohydrate metabolism; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR008979; Gal_bind_like.
DR InterPro; IPR001944; Glyco_hydro_35.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF02140; Gal_Lectin; 1.
DR Pfam; PF01301; Glyco_hydro_35; 1.
DR PRINTS; PR00742; GLYHDRLASE35.
DR ProDom; PD005612; Gal_lectin; 1.
DR PROSITE; PS01182; GLYCOSYL HYDROL F35; 1.
DR PROSITE; PS00435; PEROXIDASE 1; 1.
DR PROSITE; PS50228; SUEL LECTIN; 1.
DR Glycosidase; Hydrolase; Signal.
KW

```

```

FT SIGNAL 1 32 POTENTIAL
FT CHAIN 33 847 PUTATIVE BETA-GALACTOSIDASE.
SQ SEQUENCE 847 AA; 93658 MW; 91C13DE26A4CF4AD CRC64;

Query Match 58.0%; Score 51; DB 10; Length 847;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 539 LRAGFNKIALSIAGV 554

RESULT 14
Q8RWC1 PRELIMINARY; PRT; 847 AA.
AC Q8RWC1/2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Galactosidase, putative (EC 3.2.1.23) (Beta-galactosidase)
DE (lactase)
GN AT3G33750.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banu J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
CC GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
CC -1- SIMILARITY: BELONGS TO FAMILY 35 OF GLYCOSYL HYDROLASES.
DR EMBL; AY093197; AM13196.1; -
DR GO; GO:0009341; C:beta-galactosidase complex; IEA.
DR GO; GO:0004565; F:beta-galactosidase activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0005529; F:sugar binding; IEA.
DR GO; GO:0005975; P:response to oxidative stress; IEA.
DR InterPro; IPR008979; P:carbohydrate metabolism; IEA.
DR InterPro; IPR000922; Gal bind like.
DR InterPro; IPR001944; Glyco hydro 35.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF02140; Gal_Lectin; 1.
DR Pfam; PF01301; Glyco_Hydro_35; 1.
DR PRINTS; PR00742; GLHYDLASE35.
DR PRODOM; PD005612; Gal lectin; 1.
DR PROSITE; PS01182; GLYCOSYL HYDROL_F35; 1.
DR PROSITE; PS00435; PEROXIDASE_1; 1.
DR PROSITE; PS00228; SUEL_LECTIN; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 847 AA; 93672 MW; 0F9E12685426CSDA CRC64;

Query Match 58.0%; Score 51; DB 10; Length 847;
Best Local Similarity 62.5%; Pred. No. 10;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 539 LRAGFNKIALSIAGV 554

RESULT 15
Q7VEF7 PRELIMINARY; PRT; 610 AA.
AC Q7VEF7 (TREMELrel. 25, Created)
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE ABC-type multidrug transport system ATPase and permease components.
DE MDLB OR PRO00056.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SARG / CCMF 1375 / SS120;
RA MEDLINE=22810154; PubMed=12917486;
RA Dufrene A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RA "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RA a nearly minimal oxyphototrophic genome."
RA Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
DR EMBL; AE017161; AAP99102.1; -
DR Complete proteome.
SQ SEQUENCE 610 AA; 68701 MW; EBF39690C516DCDA CRC64;

Query Match 55.7%; Score 49; DB 16; Length 610;
Best Local Similarity 44.4%; Pred. No. 16;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSLVAV 18
DB 156 ILQIVSGFFIVSFISIAI 173

Search completed: February 26, 2004, 12:13:45
Job time : 11.1141 secs

```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 11.7931 Seconds  
(without alignments)  
455.215 Million cell updates/sec

Title: US-10-036-444-5  
Perfect score: 88  
Sequence: 1 VLLLRAGFYAVSFLSVANG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1980s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	19	4	Aae02772 Human Nkp
2	88	100.0	177	2	Aay06402 Human B-c
3	88	100.0	190	2	Aay06401 Human B-c
4	88	100.0	190	4	Aae02769 Human Nkp
5	88	100.0	201	2	Aay06403 Human B-c
6	80	90.9	19	5	Aae23900 Human Nkp
7	51	58.0	274	3	Aag23590 Arabidops
8	51	58.0	279	3	Aag23589 Arabidops
9	48	54.5	665	6	Aao16046 Carica pa
10	48	54.5	665	6	Aae32202 Papaya be
11	48	54.5	721	6	Aao16048 Carica pa
12	48	54.5	721	6	Aae32204 Papaya be
13	47	53.4	835	2	Aay44303 Tomato be
14	47	53.4	838	2	Aar82882 Tomato ex
15	47	53.4	838	3	Aay44305 Tomato be
16	46	52.3	114	3	Aag05624 Arabidops
17	46	52.3	119	3	Aag05623 Arabidops
18	46	52.3	127	3	Aag05622 Arabidops
19	46	52.3	661	3	Aag42606 Arabidops
20	46	52.3	714	3	Aag42605 Arabidops
21	46	52.3	724	3	Aay44306 Tomato be
22	46	52.3	757	3	Aag42604 Arabidops
23	46	52.3	757	5	Abb91980 Herbicida
24	45	51.1	1563	6	Abu41824 Protein e
25	44	50.0	87	7	Adc97632 E. faeciu

26	44	50.0	385	3	AAB20910
27	44	50.0	385	7	ADD15306
28	44	50.0	417	4	AG81928
29	44	50.0	425	5	ABP6312
30	44	50.0	602	5	ABP3571
31	44	50.0	602	6	ABU42891
32	43.5	49.4	383	6	ABO00770
33	43	48.9	154	6	ABO00816
34	43	48.9	166	4	ABB66413
35	43	48.9	469	4	AB76682
36	43	48.9	530	4	AB76681
37	43	48.9	530	4	AB79761
38	43	48.9	578	4	AAG90808
39	43	48.9	589	4	ABB71245
40	42	47.7	271	6	ABU35719
41	42	47.7	466	5	ABBA7399
42	41.5	47.2	588	6	AAE32079
43	41.5	47.2	589	5	AAU99329
44	41.5	47.2	589	6	AAO30994
45	41.5	47.2	589	7	ADD01392

ALIGNMENTS

RESULT 1  
AAE02772  
ID AAE02772 standard; peptide; 19 AA.  
XX  
AC AAE02772;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Human Nkp30 receptor transmembrane region sequence.  
XX  
KW Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;  
KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
KW therapy; transmembrane region.

OS Homo sapiens.  
XX  
PN WO200136630-A2.  
XX  
PD 25-MAY-2001.  
XX  
PF 15-NOV-2000; 2000WO-EP011697.  
XX  
PR 15-NOV-1999; 99CA-02288307.  
XX  
PR 15-NOV-1999; 99US-00440514.  
XX  
PA (INNA-) INNATE PHARMA SAS.  
XX  
PA (UYGE-) UNIV GENOVA.  
XX  
PI Moretta A, Bottino C, Biassoni R;  
XX  
PI MPI; 2001-329221/34.  
XX  
PT Novel compound, useful for detection and/or quantifying the presence of  
XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.  
XX  
FS Claim 1; Fig 7B; 83pp; English.  
XX  
CC The invention relates to human Nkp30 receptor and its corresponding cDNA  
CC molecule which is involved in natural cytotoxicity mediated by natural  
CC killer (NK) cells and antibodies that identify the same. Nkp30 is selectively  
CC is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively  
CC expressed on the surface of human mature NK cells. Nkp30 and its cDNA are  
CC useful for detecting and/or quantifying the presence of NK cells in a  
CC biological sample. The invention also provide kits for detecting and/or  
CC quantifying the presence of NK cells, for the selective removal of NK  
CC cells from a biological sample, for the positive and selective  
CC purification of NK cells from a biological sample and for the in vitro

stimulation of NK cell cytotoxicity. The invention further provides a pharmaceutical composition which is used as a drug for grafting enhancement, graft versus host (GVH) inhibition, stimulation of graft versus tumour (GVT) and especially graft versus leukaemia (GVL), and for the prevention, palliation and/or therapy of solid or liquid tumours, such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or microorganism, notably viral infection. Nkp30 antibodies are useful for identifying Nkp30 natural ligands and allow assessment of the level of surface Nkp30 ligand expressed on an NK-susceptible target cell and the comparison of this level to the standard physiological one. Hence Nkp30 antibodies are useful in the diagnosis of tumours or of infection. The present sequence is the transmembrane region of human Nkp30 receptor

Query Match 100.0%; Score 88; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. NO. 4.8e-08; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0;

QY 1 VLLLRAGFYAVSFSLVAVG 19  
DB 1 VLLLRAGFYAVSFSLVAVG 19

RESULT 2  
AA06402  
ID AAY06402 standard; protein; 177 AA.

AC AAY06402;  
DT 20-SEP-1999 (first entry)

DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
KW signal transduction; immunomodulator; antiinflammatory;  
KW autoimmune disease; inflammation; gene therapy; diagnosis.

OS Homo sapiens.

Key Location/Qualifiers  
Peptide 1..12 /note= "leader peptide"  
Protein 13..177 /note= "mature protein"  
Modified-site 42 /note= "N-glycosylated"  
Modified-site 68 /note= "N-glycosylated"  
Modified-site 121 /note= "N-glycosylated"  
Domain 139..162 /note= "transmembrane domain"  
Peptide 166..177 /note= "alternatively spliced C-terminal end"

WO9923867-A2.  
20-MAY-1999.  
05-NOV-1998; 98WO-US023826.  
07-NOV-1997; 97US-0064761P.  
(BIOJ ) BIOGEN INC.  
Browning J;  
WPI; 1999-418423/35.  
N-PSDB; AAX59348.

Novel B-cell myelin oligodendrocyte glycoproteins.

PS Claim 2; Page 43; 43pp; English.  
XX This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

XX SQ Sequence 177 AA;

Query Match 100.0%; Score 88; DB 2; Length 177;  
Best Local Similarity 100.0%; Pred. NO. 6.2e-07;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSLVAVG 19  
DB 139 VLLLRAGFYAVSFSLVAVG 157

RESULT 3  
AA06401  
ID AAY06401 standard; protein; 190 AA.

AC AAY06401;  
DT 20-SEP-1999 (first entry)

DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
KW signal transduction; immunomodulator; antiinflammatory;  
KW autoimmune disease; inflammation; gene therapy; diagnosis.

OS Homo sapiens.

Key Location/Qualifiers  
Peptide 1..12 /note= "leader peptide"  
Protein 13..190 /note= "mature protein"  
Modified-site 42 /note= "N-glycosylated"  
Modified-site 68 /note= "N-glycosylated"  
Modified-site 121 /note= "N-glycosylated"  
Domain 139..162 /note= "transmembrane domain"  
Peptide 166..190 /note= "alternatively spliced C-terminal end"

WO9923867-A2.  
20-MAY-1999.  
05-NOV-1998; 98WO-US023826.  
07-NOV-1997; 97US-0064761P.  
(BIOJ ) BIOGEN INC.  
Browning J;  
WPI; 1999-418423/35.

DR N-PSDB; AAX59347.  
XX Novel B-cell myelin oligodendrocyte glycoproteins.  
PT Claim 2; Page 42; 43pp; English.  
XX  
XX This sequence represents human BMOG, a novel member of the B cell myelin  
CC oligodendrocyte glycoprotein family that is expressed by germinal centre  
CC B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were  
CC identified. The protein is present primarily in the spleen, in lymph  
CC nodes and in germinal centre B cells. It may have immunoregulatory  
CC functions, and soluble or chimeric fusion proteins of BMOG may be used to  
CC regulate the immune system in autoimmune or inflammatory disease. Vectors  
CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of  
CC producing BMOG using these transformed host cells are also provided. BMOG  
CC polypeptides can be used for modulating the immune system of a subject or  
CC to inhibit signal transduction in a cell expressing BMOG by contacting it  
CC with a soluble BMOG protein. The nucleic acid can be used for gene  
CC therapy. The protein can also be used to target a toxin, imaging agent or  
CC radionuclide to a cell expressing BMOG. (All claimed)  
XX  
XX Sequence 190 AA;  
SQ  
Query Match 100.0%; Score 88; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. NO. 6.8e-07; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0;  
QY 1 VLLLRAGFYAVSFSLVAVG 19  
DB 139 VLLLRAGFYAVSFSLVAVG 157  
RESULT 4  
AAE02769  
ID AAE02769 standard; protein; 190 AA.  
XX  
AC AAE02769;  
XX  
DT 06-AUG-2001 (first entry)  
XX  
DE Human Nkp30 receptor.  
XX  
XX Human, Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;  
KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
KW therapy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..18  
FT /label= Signal\_peptide  
FT Protein 19..190  
FT /label= Mature\_Nkp30\_receptor\_protein  
FT Region 19..138  
FT /label= Extracellular\_region  
FT Modified-site 42  
FT /note= "Forms an immunoglobulin (Ig) V-like domain"  
FT Modified-site 121  
FT /note= "N-glycosylation site"  
FT Modified-site 139..157  
FT /note= "N-glycosylation site"  
FT Region 139..157  
FT /label= Transmembrane\_region  
FT Region 158..190  
FT /label= Intracellular\_region  
XX  
XX WO200136630-A2.  
XX  
XX 25-MAY-2001.  
PD  
XX 15-NOV-2000; 2000WO-EP011697.  
XX  
XX 15-NOV-1999; 99CA-02288307.  
PR

PR 15-NOV-1999; 99US-00440514.  
XX (INNA-) INNATE PHARMA SAS.  
PA (UYGE-) UNIV GENOVA.  
XX  
XX Moretta A, Bottino C, Biassoni R;  
XX  
XX WPI; 2001-329221/34.  
DR N-PSDB; AAD06564.  
XX  
XX Novel compound, useful for detection and/or quantifying the presence of  
PT NK cells, comprises the amino acid sequences of the Nkp30 molecule.  
XX  
XX Claim 1; Fig 7B; 83pp; English.  
XX  
XX The invention relates to human Nkp30 receptor and its corresponding cDNA  
CC molecule which is involved in natural cytotoxicity mediated by natural  
CC killer (NK) cells and antibodies that identify the same. Nkp30 receptor  
CC is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively  
CC expressed on the surface of human mature NK cells. Nkp30 and its cDNA are  
CC useful for detecting and/or quantifying the presence of NK cells in a  
CC biological sample. The invention also provides kits for detecting and/or  
CC quantifying the presence of NK cells, for the selective removal of NK  
CC cells from a biological sample, for the positive and selective  
CC purification of NK cells from a biological sample and for the in vitro  
CC stimulation of NK cell cytotoxicity. The invention further provides a  
CC pharmaceutical composition which is used as a drug for grafting  
CC enhancement, graft versus host (GvH) inhibition, stimulation of graft  
CC versus tumour (GvT) and especially graft versus leukaemia (GvL), and for  
CC the prevention, palliation and/or therapy of solid or liquid tumours,  
CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or  
CC microorganism, notably viral infection. Nkp30 antibodies are useful for  
CC identifying Nkp30 natural ligands and allow assessment of the level of  
CC surface Nkp30 ligand expressed on an NK-susceptible target cell and the  
CC comparison of this level to the standard physiological one. Hence Nkp30  
CC antibodies are useful in the diagnosis of tumours or of infection. The  
CC present sequence is human Nkp30 receptor  
XX  
XX Sequence 190 AA;  
SQ  
Query Match 100.0%; Score 88; DB 4; Length 190;  
Best Local Similarity 100.0%; Pred. NO. 6.8e-07; Indels 0; Gaps 0;  
Matches 19; Conservative 0; Mismatches 0;  
QY 1 VLLLRAGFYAVSFSLVAVG 19  
DB 139 VLLLRAGFYAVSFSLVAVG 157  
RESULT 5  
AAAY06403  
ID AAAY06403 standard; protein; 201 AA.  
XX  
AC AAAY06403;  
XX  
XX 20-SEP-1999 (first entry)  
XX  
XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
XX  
XX BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
KW signal transduction; immunomodulator; antiinflammatory;  
KW autoimmune disease; inflammation; gene therapy; diagnosis.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..12  
FT /note= "leader peptide"  
FT Protein 13..201  
FT /note= "mature protein"  
FT Modified-site 42  
FT /note= "N-glycosylated"  
FT Modified-site 68

FT Modified-site /note= "N-glycosylated"  
 FT 121  
 FT Domain /note= "N-glycosylated"  
 FT 139..162  
 FT /note= "transmembrane domain"  
 FT 166..201  
 FT Peptide /note= "alternatively spliced C-terminal end"  
 FT XX  
 FT WO9923867-A2.  
 FT XX  
 FT 20-MAY-1999.  
 FT XX  
 FT 05-NOV-1998; 98WO-US023826.  
 FT XX  
 FT 07-NOV-1997; 97US-0064761P.  
 FT XX  
 FT (BIOJ ) BIOGEN INC.  
 FT XX  
 FT Browning J;  
 FT XX  
 FT WPI; 1999-418423/35.  
 FT DR N-PSDB; AAX59349.  
 FT XX  
 FT Novel B-cell myelin oligodendrocyte glycoproteins.  
 FT XX  
 FT Claim 2; Page 43; 43pp; English.  
 FT PS  
 FT This sequence represents human BMOG, a novel member of the B cell myelin  
 CC oligodendrocyte glycoprotein family that is expressed by germinal centre  
 CC B cells. 3 C-terminal splice variants (see AA06401-03) of BMOG were  
 CC identified. The protein is present primarily in the spleen, in lymph  
 CC nodes and in germinal centre B cells. It may have immunoregulatory  
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to  
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors  
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of  
 CC producing BMOG using these transformed host cells are also provided. BMOG  
 CC polypeptides can be used for modulating the immune system of a subject or  
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it  
 CC with a soluble BMOG protein. The nucleic acid can be used for gene  
 CC therapy. The protein can also be used to target a toxin, imaging agent or  
 CC radionuclide to a cell expressing BMOG. (All claimed)  
 CC XX  
 CC Sequence 201 AA;  
 CC SQ  
 Query Match 100.0%; Score 88; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-07;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 VLLLRAGFYAVSFSLVAVG 19  
 DB 139 VLLLRAGFYAVSFSLVAVG 157  
 RESULT 6  
 AAEE23900  
 ID AAEE23900 standard; peptide; 19 AA.  
 XX  
 AC AAEE23900;  
 XX  
 DT 10-SEP-2002 (first entry)  
 XX  
 DE Human Nkp30 receptor peptide.  
 XX  
 KW KAR-associated protein; KARAP-transduced immune signal; dendritic cell;  
 KW antigen presentation; contact sensitivity; multiple sclerosis;  
 KW neuroprotective; human; Nkp30 receptor peptide.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200224940-A2.  
 XX  
 XX 28-MAR-2002.  
 PD  
 XX

PF 20-SEP-2001; 2001WO-EP011492.  
 XX  
 PR 20-SEP-2000; 2000US-0234161P.  
 XX  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 XX  
 XX Vivier E, Vely F, Tomasello E;  
 XX  
 XX WPI; 2002-454420/48.  
 DR  
 XX  
 XX Identifying KAR-associated protein-transduced immune signal inhibitor,  
 PT comprises using cells co-expressing functional KARAP, and engineered  
 PT cells and animals that over-express functional KARAP or bear non-  
 PT functional KARAP.  
 XX  
 XX Example 4; Page 45; 89pp; English.  
 PS  
 XX  
 CC The present invention relates to a novel method for identifying compounds  
 CC capable of inhibiting KAR-associated protein (KARAP)-transduced immune  
 CC signals. The method involves using functional and non-functional KARAP,  
 CC cells co-expressing functional KARAP, functional receptors transducing  
 CC their signal by zeta, gamma or epsilon and engineered cells and animals  
 CC over-expressing functional KARAP or bearing non-functional KARAP. The  
 CC method is useful for identifying compounds capable of inhibiting KARAP-  
 CC transduced immune signals. The KARAP-inhibiting compounds are useful for  
 CC impairing the development and maturation of dendritic cells, for  
 CC inhibiting the antigen presentation of dendritic cells, by synthesis  
 CC inhibition or through inhibition of the migration of dendritic cells, for  
 CC making drugs intended for inhibiting dendritic cell development or  
 CC maturation, for preparing drugs for the treatment, prevention, palliation  
 CC of immune response, where the activation of KAR has to be inhibited and  
 CC for the treatment of contact sensitivity or multiple sclerosis. The  
 CC present sequence is human Nkp30 receptor peptide which associate with CD3  
 CC zeta and FcR gamma. This sequence is used in the exemplification of the  
 CC invention  
 CC XX  
 CC Sequence 19 AA;  
 CC SQ  
 Query Match 90.9%; Score 80; DB 5; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 3 LLRAGFYAVSFSLVAVG 19  
 DB 1 LLRAGFYAVSFSLVAVG 17  
 RESULT 7  
 AAG23590  
 ID AAG23590 standard; protein; 274 AA.  
 XX  
 XX AAG23590;  
 AC  
 XX  
 XX 17-OCT-2000 (first entry)  
 DT  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 26954.  
 XX  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 XX Arabidopsis thaliana.  
 OS  
 XX BP1033405-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX 25-FEB-2000; 2000EP-00301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825P.  
 XX  
 PR 05-MAR-1999; 99US-0123180P.  
 PR  
 PR 09-MAR-1999; 99US-0123548P.  
 PR  
 PR 23-MAR-1999; 99US-0125788P.  
 PR



PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 28-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 05-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.

PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149733P.  
 PR 20-AUG-1999; 99US-0149923P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157533P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.

```

PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159337P.
PR 14-OCT-1999; 99US-0159638P.
PR 14-OCT-1999; 99US-0159683P.
PR 18-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 58.08; Score 51; DB 3; Length 274;
Best Local Similarity 62.5%; Pred. No. 2.1;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 4 LRAGFYAVSFSLVAVG 19
Db 30 LRAGFNKIALISIAVG 45

RESULT 8
AAG23589
ID AAG23589 standard; protein; 279 AA.
AC AAG23589;
XX
XX
DT 17-OCT-2000 (first entry)
DE
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137744P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139432P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142053P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.

```

PR 20-JUL-1999;	99US-01443352P.	PR 14-OCT-1999;	99US-0159331P.	Query Match	58.0%;	Score 51;	DB 3;	Length 279;
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.	Best Local Similarity	62.5%;	Pred. No. 2.2;		
PR 20-JUL-1999;	99US-0144884P.	PR 14-OCT-1999;	99US-0159638P.	Matches	10;	Conservative	3;	Mismatches
PR 21-JUL-1999;	99US-0144814P.	PR 18-OCT-1999;	99US-0159584P.				3;	Indels
PR 21-JUL-1999;	99US-0145086P.	PR 21-OCT-1999;	99US-0160741P.				0;	Gaps
PR 21-JUL-1999;	99US-0145088P.	PR 21-OCT-1999;	99US-0160767P.				0;	Gaps
PR 22-JUL-1999;	99US-0145085P.	PR 21-OCT-1999;	99US-0160768P.					
PR 22-JUL-1999;	99US-0145087P.	PR 21-OCT-1999;	99US-0160770P.					
PR 22-JUL-1999;	99US-0145089P.	PR 21-OCT-1999;	99US-0160814P.					
PR 22-JUL-1999;	99US-0145192P.	PR 21-OCT-1999;	99US-0160815P.					
PR 23-JUL-1999;	99US-0145145P.	PR 22-OCT-1999;	99US-0160980P.					
PR 23-JUL-1999;	99US-0145218P.	PR 22-OCT-1999;	99US-0160981P.					
PR 23-JUL-1999;	99US-0145224P.	PR 22-OCT-1999;	99US-0160989P.					
PR 23-JUL-1999;	99US-0145276P.	PR 25-OCT-1999;	99US-0161404P.					
PR 23-JUL-1999;	99US-0145313P.	PR 25-OCT-1999;	99US-0161405P.					
PR 27-JUL-1999;	99US-0145918P.	PR 25-OCT-1999;	99US-0161406P.					
PR 27-JUL-1999;	99US-0145919P.	PR 26-OCT-1999;	99US-0161359P.					
PR 28-JUL-1999;	99US-0145951P.	PR 26-OCT-1999;	99US-0161360P.					
PR 02-AUG-1999;	99US-0146386P.	PR 26-OCT-1999;	99US-0161361P.					
PR 02-AUG-1999;	99US-0146388P.	PR 28-OCT-1999;	99US-0161920P.					
PR 02-AUG-1999;	99US-0146389P.	PR 28-OCT-1999;	99US-0161992P.					
PR 03-AUG-1999;	99US-0147038P.	PR 28-OCT-1999;	99US-0161993P.					
PR 04-AUG-1999;	99US-0147204P.	PR 29-OCT-1999;	99US-0162142P.					
PR 04-AUG-1999;	99US-0147302P.							
PR 05-AUG-1999;	99US-0147192P.							
PR 05-AUG-1999;	99US-0147260P.							
PR 06-AUG-1999;	99US-0147303P.							
PR 06-AUG-1999;	99US-0147416P.							
PR 09-AUG-1999;	99US-0147493P.							
PR 09-AUG-1999;	99US-0147935P.							
PR 10-AUG-1999;	99US-0148171P.							
PR 11-AUG-1999;	99US-0148319P.							
PR 12-AUG-1999;	99US-0148341P.							
PR 13-AUG-1999;	99US-0148565P.							
PR 13-AUG-1999;	99US-0148684P.							
PR 16-AUG-1999;	99US-0149388P.							
PR 17-AUG-1999;	99US-0149175P.							
PR 18-AUG-1999;	99US-0149426P.							
PR 20-AUG-1999;	99US-0149723P.							
PR 20-AUG-1999;	99US-0149929P.							
PR 23-AUG-1999;	99US-0149902P.							
PR 23-AUG-1999;	99US-0149930P.							
PR 25-AUG-1999;	99US-0150566P.							
PR 26-AUG-1999;	99US-0150884P.							
PR 27-AUG-1999;	99US-0151065P.							
PR 27-AUG-1999;	99US-0151066P.							
PR 30-AUG-1999;	99US-0151080P.							
PR 31-AUG-1999;	99US-0151303P.							
PR 01-SEP-1999;	99US-0151930P.							
PR 07-SEP-1999;	99US-0152363P.							
PR 10-SEP-1999;	99US-0153070P.							
PR 13-SEP-1999;	99US-0153758P.							
PR 15-SEP-1999;	99US-0154018P.							
PR 16-SEP-1999;	99US-0154039P.							
PR 20-SEP-1999;	99US-0154779P.							
PR 22-SEP-1999;	99US-0155139P.							
PR 23-SEP-1999;	99US-0155486P.							
PR 24-SEP-1999;	99US-0155659P.							
PR 28-SEP-1999;	99US-0156458P.							
PR 29-SEP-1999;	99US-0156596P.							
PR 04-OCT-1999;	99US-0157117P.							
PR 05-OCT-1999;	99US-0157753P.							
PR 06-OCT-1999;	99US-0157865P.							
PR 08-OCT-1999;	99US-0158029P.							
PR 12-OCT-1999;	99US-0158369P.							
PR 13-OCT-1999;	99US-0159293P.							
PR 13-OCT-1999;	99US-0159294P.							
PR 13-OCT-1999;	99US-0159299P.							
PR 14-OCT-1999;	99US-0159325P.							
PR 14-OCT-1999;	99US-0159330P.							
QY	4 LRAGFYAVSPLSVAVG 19	QY	4 LRAGFYAVSPLSVAVG 19	RESULT 9				
DB	35 LRAGFNKIALISVAVG 50	DB	35 LRAGFNKIALISVAVG 50	AAO16046				
				ID AAO16046 standard; protein; 665 AA.				
				XX AAO16046;				
				XX 27-FEB-2003 (first entry)				
				XX Carica papaya beta-galactosidase (Beta-gal-45).				
				XX Enzyme; papaya; beta-galactosidase; pectinmethylesterase;				
				XX Polygalacturonase; fruit ripening.				
				XX Carica papaya.				
				XX WO200283924-A2.				
				XX 24-OCT-2002.				
				XX 11-APR-2002; 2002WO-US011804.				
				XX 11-APR-2001; 2001US-0283008P.				
				XX (CORR ) CORNELL RES FOUND INC.				
				XX (PAIS/) PAIS M S S.				
				XX Pais MSS, Gonsalves D, Balde A;				
				XX WPI; 2003-075556/07.				
				XX N-PSDB; AAL51072.				
				XX New isolated nucleic acid molecule encoding beta-galactosidase protein,				
				XX useful for promoting or delaying papaya fruit ripening.				
				XX Claim 3; Page 9-13; 84pp; English.				
				XX The invention comprises the amino acid and coding sequence of Carica				
				XX papaya beta-galactosidase proteins which control papaya fruit ripening.				
				XX The invention also comprises the amino acid and coding sequence of Carica				
				XX papaya pectinmethylesterase and polygalacturonase proteins. The DNA and				

CC protein sequences of the invention are useful for promoting or delaying  
 CC papaya fruit ripening. The present amino acid sequence represents a  
 CC Carica papaya fruit ripening-related protein of the invention

XX SQ Sequence 665 AA;  
 Query Match 54.5%; Score 48; DB 6; Length 665;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19  
 ||||| :|||:  
 Db 383 LRAGINKISILSIIVG 398

RESULT 10  
 AAE32202  
 ID AAE32202 standard; protein; 665 AA.  
 AC AAE32202;  
 XX AAE32202;  
 XX 24-MAR-2003 (first entry)  
 DT 24-MAR-2003 (first entry)  
 DE Papaya beta-galactosidase (beta-Gal).45.  
 XX Fruit ripening; papaya ringspot virus coat protein; transgenic plant;  
 XX transgenic; PRSV; CP; beta-galactosidase; beta-Gal.45; enzyme.  
 KW Carica papaya.  
 OS Carica papaya.

XX Key Location/Qualifiers  
 FH Misc-difference 143  
 FT /label= "Unknown"  
 FT /note= "Encoded by TGN"  
 XX WO200282889-A1.  
 PN 24-OCT-2002.  
 XX 24-OCT-2002.  
 PD 11-APR-2002; 2002WO-US011803.  
 PF 11-APR-2001; 2001US-0283022P.  
 PR (CORR ) CORNELL RES FOUND INC.  
 XX (PAIS/) PAIS M S S.  
 PA Pais MSS, Gonsalves D, Balde A, Chiang C;  
 PI WPI; 2003-075493/07.  
 XX N-PSDB; AAD49832.  
 DR New DNA construct, useful for controlling the ripening of papaya fruit  
 PT and conferring resistance to papaya ringspot virus coat in transgenic  
 PT plants.  
 XX Disclosure; Page 75-77; 121pp; English.

CC The invention relates to a DNA construct comprising a first DNA molecule  
 CC encoding a protein that controls papaya fruit ripening and a second DNA  
 CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The  
 CC DNA construct is useful for controlling the ripening of papaya fruit and  
 CC conferring resistance to PRSV coat in transgenic plants. The present  
 CC sequence is papaya beta-galactosidase (beta-Gal).45  
 XX Sequence 665 AA;  
 Query Match 54.5%; Score 48; DB 6; Length 665;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19  
 ||||| :|||:  
 Db 383 LRAGINKISILSIIVG 398

RESULT 11  
 AAO16048  
 ID AAO16048 standard; protein; 721 AA.  
 AC AAO16048;  
 XX 27-FEB-2003 (first entry)  
 DT 27-FEB-2003 (first entry)  
 DE Carica papaya beta-galactosidase (beta-gal-41).  
 XX Enzyme; papaya; beta-galactosidase; pectinmethylesterase;  
 KW polygalacturonase; fruit ripening.  
 XX Carica papaya.  
 OS Carica papaya.  
 XX WO200283924-A2.  
 PN 24-OCT-2002.  
 PD 11-APR-2002; 2002WO-US011804.  
 PF 11-APR-2001; 2001US-0283008P.  
 PR (CORR ) CORNELL RES FOUND INC.  
 XX (PAIS/) PAIS M S S.  
 PA Pais MSS, Gonsalves D, Balde A;  
 PI WPI; 2003-075556/07.  
 XX N-PSDB; AAL51074.  
 DR New isolated nucleic acid molecule encoding beta-galactosidase protein,  
 PT useful for promoting or delaying papaya fruit ripening.  
 PT Claim 17; Page 15-19; 84pp; English.  
 XX The invention comprises the amino acid and coding sequence of Carica  
 XX papaya beta-galactosidase proteins which control papaya fruit ripening.  
 CC The invention also comprises the amino acid and coding sequence of Carica  
 CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and  
 CC protein sequences of the invention are useful for promoting or delaying  
 CC papaya fruit ripening. The present amino acid sequence represents a  
 CC Carica papaya fruit ripening-related protein of the invention

XX SQ Sequence 721 AA;  
 Query Match 54.5%; Score 48; DB 6; Length 721;  
 Best Local Similarity 68.8%; Pred. No. 21;  
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19  
 ||||| :|||:  
 Db 526 LRAGVNRKVSLSIIVG 541

RESULT 12  
 AAE32204  
 ID AAE32204 standard; protein; 721 AA.  
 AC AAE32204;  
 XX 24-MAR-2003 (first entry)  
 DT 24-MAR-2003 (first entry)  
 DE Papaya beta-galactosidase (beta-Gal).41.  
 XX Fruit ripening; papaya ringspot virus coat protein; transgenic plant;  
 KW transgenic; PRSV; CP; beta-galactosidase; beta-Gal.41; enzyme.  
 XX Carica papaya.  
 OS Carica papaya.  
 XX WO200282889-A1.  
 PN 24-OCT-2002.  
 PD 11-APR-2002; 2002WO-US011803.  
 PF 11-APR-2001; 2001US-0283022P.  
 PR (CORR ) CORNELL RES FOUND INC.  
 XX (PAIS/) PAIS M S S.  
 PA Pais MSS, Gonsalves D, Balde A, Chiang C;  
 PI WPI; 2003-075493/07.  
 XX N-PSDB; AAD49832.  
 DR New DNA construct, useful for controlling the ripening of papaya fruit  
 PT and conferring resistance to papaya ringspot virus coat in transgenic  
 PT plants.  
 XX Disclosure; Page 75-77; 121pp; English.

CC The invention relates to a DNA construct comprising a first DNA molecule  
 CC encoding a protein that controls papaya fruit ripening and a second DNA  
 CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The  
 CC DNA construct is useful for controlling the ripening of papaya fruit and  
 CC conferring resistance to PRSV coat in transgenic plants. The present  
 CC sequence is papaya beta-galactosidase (beta-Gal).45  
 XX Sequence 665 AA;  
 Query Match 54.5%; Score 48; DB 6; Length 665;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19  
 ||||| :|||:  
 Db 383 LRAGINKISILSIIVG 398

RESULT 10  
 AAE32202  
 ID AAE32202 standard; protein; 665 AA.  
 AC AAE32202;  
 XX AAE32202;  
 XX 24-MAR-2003 (first entry)  
 DT 24-MAR-2003 (first entry)  
 DE Papaya beta-galactosidase (beta-Gal).45.  
 XX Fruit ripening; papaya ringspot virus coat protein; transgenic plant;  
 XX transgenic; PRSV; CP; beta-galactosidase; beta-Gal.45; enzyme.  
 KW Carica papaya.  
 OS Carica papaya.  
 XX WO200282889-A1.  
 PN 24-OCT-2002.  
 PD 11-APR-2002; 2002WO-US011803.  
 PF 11-APR-2001; 2001US-0283022P.  
 PR (CORR ) CORNELL RES FOUND INC.  
 XX (PAIS/) PAIS M S S.  
 PA Pais MSS, Gonsalves D, Balde A, Chiang C;  
 PI WPI; 2003-075493/07.  
 XX N-PSDB; AAD49832.  
 DR New DNA construct, useful for controlling the ripening of papaya fruit  
 PT and conferring resistance to papaya ringspot virus coat in transgenic  
 PT plants.  
 XX Disclosure; Page 75-77; 121pp; English.

CC The invention relates to a DNA construct comprising a first DNA molecule  
 CC encoding a protein that controls papaya fruit ripening and a second DNA  
 CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The  
 CC DNA construct is useful for controlling the ripening of papaya fruit and  
 CC conferring resistance to PRSV coat in transgenic plants. The present  
 CC sequence is papaya beta-galactosidase (beta-Gal).45  
 XX Sequence 665 AA;  
 Query Match 54.5%; Score 48; DB 6; Length 665;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19  
 ||||| :|||:  
 Db 383 LRAGINKISILSIIVG 398

RESULT 11  
 AAO16048  
 ID AAO16048 standard; protein; 721 AA.  
 AC AAO16048;  
 XX 27-FEB-2003 (first entry)  
 DT 27-FEB-2003 (first entry)  
 DE Carica papaya beta-galactosidase (beta-gal-41).  
 XX Enzyme; papaya; beta-galactosidase; pectinmethylesterase;  
 KW polygalacturonase; fruit ripening.  
 XX Carica papaya.  
 OS Carica papaya.  
 XX WO200283924-A2.  
 PN 24-OCT-2002.  
 PD 11-APR-2002; 2002WO-US011804.  
 PF 11-APR-2001; 2001US-0283008P.  
 PR (CORR ) CORNELL RES FOUND INC.  
 XX (PAIS/) PAIS M S S.  
 PA Pais MSS, Gonsalves D, Balde A;  
 PI WPI; 2003-075556/07.  
 XX N-PSDB; AAL51074.  
 DR New isolated nucleic acid molecule encoding beta-galactosidase protein,  
 PT useful for promoting or delaying papaya fruit ripening.  
 PT Claim 17; Page 15-19; 84pp; English.  
 XX The invention comprises the amino acid and coding sequence of Carica  
 XX papaya beta-galactosidase proteins which control papaya fruit ripening.  
 CC The invention also comprises the amino acid and coding sequence of Carica  
 CC papaya pectinmethylesterase and polygalacturonase proteins. The DNA and  
 CC protein sequences of the invention are useful for promoting or delaying  
 CC papaya fruit ripening. The present amino acid sequence represents a  
 CC Carica papaya fruit ripening-related protein of the invention

XX SQ Sequence 721 AA;  
 Query Match 54.5%; Score 48; DB 6; Length 721;  
 Best Local Similarity 68.8%; Pred. No. 21;  
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19  
 ||||| :|||:  
 Db 526 LRAGVNRKVSLSIIVG 541

XX PD 24-OCT-2002.  
XX PF 11-APR-2002; 2002WO-US011803.  
XX PR 11-APR-2001; 2001US-0283022P.  
XX PA (CORR ) CORNELL RES FOUND INC.  
XX PA (PAIS/) PAIS M S S.  
XX PI Pais MSS, Gonsalves D, Balde A, Chiang C;  
XX PF WPI; 2003-075493/07.  
XX DR N-PSDB; AAD49834.  
XX XX  
XX PF New DNA construct, useful for controlling the ripening of papaya fruit  
XX PT and conferring resistance to papaya ringspot virus coat in transgenic  
XX PT plants.  
XX PS Disclosure; Page 80-83; 121pp; English.  
XX CC The invention relates to a DNA construct comprising a first DNA molecule  
XX CC encoding a protein that controls papaya fruit ripening and a second DNA  
XX CC molecule encoding a papaya ringspot virus (PRSV) coat protein (CP). The  
XX CC DNA construct is useful for controlling the ripening of papaya fruit and  
XX CC conferring resistance to PRSV coat in transgenic plants. The present  
XX CC sequence is papaya beta-galactosidase (beta-Gal).<sup>41</sup>  
XX SQ Sequence 721 AA;  
XX  
XX Query Match 54.5%; Score 48; DB 6; Length 721;  
XX Best Local Similarity 68.8%; Pred. No. 21;  
XX Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
XX  
XX QY 4 LRAGFYAVSFSLVAVG 19  
XX ||||| :||:||||  
XX DB 526 LRAGVNVKLSLSIAVG 541  
XX  
XX RESULT 13  
XX AAY44303  
XX ID AAY44303 standard; protein; 835 AA.  
XX AC AAY44303;  
XX XX 29-FEB-2000 (first entry)  
XX DT Tomato beta galactosidase-1.  
XX DE Tomato beta galactosidase-1; TBG; Rutgers tomato plant; pectin;  
XX KW fruit softening; beta galactosidase II protein; biofilm;  
XX KW transgenic plant; protoplast isolation.  
XX XX Lycopersicon esculentum.  
XX OS Lycopersicon esculentum.  
XX FH Key Location/Qualifiers  
XX FT Peptide 1..24  
XX FT /label= Signal\_peptide  
XX FT Protein 25..835  
XX FT /label= beta-galactosidase-1  
XX PN WO9964564-A1.  
XX PD 16-DEC-1999.  
XX PF 08-JUN-1999; 99WO-US012697.  
XX XX 09-JUN-1999; 98US-0088805P.  
XX XX (USDA ) US DEPT OF AGRICULTURE.  
XX PI Gross KC, Smith DL;  
XX

DR WPI; 2000-097532/08.  
DR N-PSDB; AAZ29338.  
XX PF New beta-galactosidases, used to prepare transgenic plants with altered  
XX FT fruit ripening.  
XX PS Claim 1; Fig 2; 85pp; English.  
XX XX  
XX CC The present sequence is tomato beta galactosidase-1 (TBG-1) encoded by a  
XX CC cDNA derived from breaker, turning and pink fruit pericarp from 'Rutgers'  
XX CC tomato plants. This hydrolyses terminal non-reducing beta-D-galactosyl  
XX CC residues from beta-D-galactosides leading to loss of tissue integrity and  
XX CC fruit softening. This is used for modifying cell wall metabolism and  
XX CC controlling ripening of fruit by altering activity of beta galactosidase  
XX CC II protein. Pectin with reduced galactosyl content is produced for use in  
XX CC biofilms or solutions. Transgenic plants with altered fruit ripening are  
XX CC produced by introducing DNA constructs comprising TBG cDNA. TBG forms a  
XX CC component of an enzyme mixture used to isolate protoplasts  
XX SQ Sequence 835 AA;  
XX  
XX Query Match 53.4%; Score 47; DB 3; Length 835;  
XX Best Local Similarity 62.5%; Pred. No. 37;  
XX Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
XX  
XX QY 4 LRAGFYAVSFSLVAVG 19  
XX ||||| :||:||||  
XX DB 527 LRAGVNVKLSLSIAVG 542  
XX  
XX RESULT 14  
XX AAR82882  
XX ID AAR82882 standard; protein; 838 AA.  
XX AC AAR82882;  
XX XX 10-FEB-1996 (first entry)  
XX DT Tomato exo-(1-4)-beta-D-galactanase enzyme.  
XX DE Exo-(1-4)-beta-D-galactanase; enzyme; transgenic plant; crop improvement.  
XX KW Exo-(1-4)-beta-D-galactanase; enzyme; transgenic plant; crop improvement.  
XX XX Lycopersicon esculentum.  
XX OS Lycopersicon esculentum.  
XX PN WO9523228-A1.  
XX PD 31-AUG-1995.  
XX PF 23-FEB-1995; 95WO-GB000372.  
XX PR 23-FEB-1994; 94GB-00003423.  
XX XX (UNIL ) UNILEVER PLC.  
XX PA (UNIL ) UNILEVER NV.  
XX XX Chengappa S, Heilyer SA, De Silva J, Reid JSG;  
XX XX WPI; 1995-311537/40.  
XX DR N-PSDB; AAT01015.  
XX XX  
XX PT Lupin and tomato exo-galactanase DNA and protein - useful for the  
XX PT alteration of plant characteristic(s), e.g. texture, growth, ripening.  
XX PS Disclosure; Page 38; 67pp; English.  
XX XX  
XX CC This enzyme may be expressed in a transgenic plant in order to alter  
XX CC characteristics of e.g. growth, texture or ripening of the plant or plant  
XX CC parts. This protein may also be expressed in a recombinant host for the  
XX CC production of the enzyme which may be used for the modification,  
XX CC degradation or liquefaction of plant materials in order to affect  
XX CC mechanical properties relating to eating texture, particle sizes of, e.g.  
XX CC fruit or vegetable juices, or extractability of colors, flavors or  
XX CC vitamins



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 6.60212 Seconds  
(without alignments)  
607.670 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 88

Sequence: 1 VLLLRAGFAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	100.0	19	13	US-10-036-444-5
2	88	100.0	190	13	Sequence 5, Appli
3	48	54.5	665	14	Sequence 2, Appli
4	48	54.5	665	15	Sequence 2, Appli
5	48	54.5	721	14	Sequence 6, Appli
6	48	54.5	721	15	Sequence 6, Appli
7	44.5	50.6	473	15	Sequence 10510, A
8	43	48.9	536	10	Sequence 42, Appli
9	43	48.9	578	9	Sequence 4562, Ap
10	41.5	47.2	589	9	Sequence 2, Appli
11	41.5	47.2	589	14	Sequence 2, Appli
12	41.5	47.2	850	9	Sequence 3, Appli
13	41	46.6	297	9	Sequence 6265, Ap
14	41	46.6	320	14	Sequence 9631, Ap
15	41	46.6	333	14	Sequence 441, App

16	41	46.6	405	10	US-09-932-227-60
17	41	46.6	405	14	US-10-183-708-60
18	41	46.6	421	14	US-10-156-761-12314
19	40	45.5	97	9	US-09-738-626-3787
20	40	45.5	99	9	US-09-864-761-42102
21	40	45.5	220	15	US-10-369-493-12842
22	40	45.5	257	14	US-10-156-761-13244
23	40	45.5	278	14	US-10-029-386-32250
24	40	45.5	351	10	US-09-882-227-434
25	40	45.5	477	14	US-10-245-537A-4
26	40	45.5	551	13	US-10-216-355-4
27	39	44.3	55	11	US-09-864-408A-2362
28	39	44.3	69	14	US-10-348-713-3
29	39	44.3	90	14	US-10-106-698-4634
30	39	44.3	115	14	US-10-156-761-9869
31	39	44.3	268	9	US-09-949-192-53
32	39	44.3	376	15	US-10-622-896-2
33	39	44.3	381	15	US-10-622-896-4
34	39	44.3	394	14	US-10-081-816-40
35	39	44.3	397	15	US-10-447-328-60
36	39	44.3	479	15	US-10-369-493-20600
37	39	44.3	1263	10	US-09-882-694-11
38	39	44.3	1548	15	US-10-369-493-6347
39	38.5	43.8	560	9	US-09-931-212A-4
40	38.5	43.8	560	9	US-09-915-181A-5
41	38.5	43.8	560	10	US-09-965-522-4
42	38.5	43.8	560	15	US-10-314-790-5
43	38.5	43.8	578	9	US-09-740-041-4
44	38.5	43.8	578	14	US-10-389-967-4
45	38.5	43.8	582	9	US-09-915-181A-4

## ALIGNMENTS

### RESULT 1

US-10-036-444-5  
; Sequence 5, Application US/10036444  
; Publication No. US20020142445A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; TITLE OF INVENTION: "NO. US20020142445A1 triggering receptor involved in natural  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and  
; TITLE OF INVENTION: antibodies that identify the same"  
; FILE REFERENCE: SEQ-FR-1060  
; CURRENT APPLICATION NUMBER: US/10/036.444  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: 09/440,514  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: 09/456,199  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Human NK cell  
US-10-036-444-5

Query Match 100.0%; Score 88; DB 13; Length 19;  
Best Local Similarity 100.0%; Pred. No. 5.7e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLLLRAGFAVSFLSVAVG 19  
Db 1 VLLLRAGFAVSFLSVAVG 19

### RESULT 2

US-10-036-444-2  
; Sequence 2, Application US/10036444  
; Publication No. US20020142445A1

```

; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036,444
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Human NK cell
; US-10-036-444-2

```

```

Query Match 100.0%; Score 88; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. No. 7.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VLLLRAGFYAVSFLSVAVG 19
DB 139 VLLLRAGFYAVSFLSVAVG 157

```

```

RESULT 3
US-10-121-393-2
; Sequence 2, Application US/10121393
; Publication No. US20030115633A1
; GENERAL INFORMATION:
; APPLICANT: Pais, Marie Salome Soares
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES RELATING TO PAPAYA RIPENING
; FILE REFERENCE: 19603/3001
; CURRENT APPLICATION NUMBER: US/10/121.393
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/283,008
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Carica papaya
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (143)
; OTHER INFORMATION: xaa at position 143 in this sequence is any amino
; OTHER INFORMATION: acid
; US-10-121-393-2

```

```

Query Match 54.5%; Score 48; DB 14; Length 665;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 LRAGFYAVSFLSVAVG 19
DB 383 LRAGINKISILSIIVG 398

```

```

RESULT 4
US-10-121-539-2
; Sequence 2, Application US/10121539
; Publication No. US20030204869A1
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje

```

```

; APPLICANT: Pais, Maria Salome Soares
; APPLICANT: Chiang, Chu-Hui
; TITLE OF INVENTION: A METHOD TO CONTROL THE RIPENING OF PAPAYA FRUIT AND
; TITLE OF INVENTION: CONFER DISEASE RESISTANCE TO PAPAYA PLANTS
; FILE REFERENCE: 19603/3601
; CURRENT APPLICATION NUMBER: US/10/121.539
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/283,022
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Carica papaya
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (143)
; OTHER INFORMATION: xaa at position 143 in this sequence is any amino
; OTHER INFORMATION: acid
; US-10-121-539-2

```

```

Query Match 54.5%; Score 48; DB 15; Length 665;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 LRAGFYAVSFLSVAVG 19
DB 383 LRAGINKISILSIIVG 398

```

```

RESULT 5
US-10-121-393-6
; Sequence 6, Application US/10121393
; Publication No. US20030115633A1
; GENERAL INFORMATION:
; APPLICANT: Pais, Marie Salome Soares
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES RELATING TO PAPAYA RIPENING
; FILE REFERENCE: 19603/3001
; CURRENT APPLICATION NUMBER: US/10/121.393
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/283,008
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 721
; TYPE: PRT
; ORGANISM: Carica papaya
; US-10-121-393-6

```

```

Query Match 54.5%; Score 48; DB 14; Length 721;
Best Local Similarity 68.8%; Pred. No. 14;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 LRAGFYAVSFLSVAVG 19
DB 526 LRAGINKISILSIIVG 541

```

```

RESULT 6
US-10-121-539-6
; Sequence 6, Application US/10121539
; Publication No. US20030204869A1
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Balde, Aladjje
; APPLICANT: Pais, Maria Salome Soares
; APPLICANT: Chiang, Chu-Hui
; TITLE OF INVENTION: A METHOD TO CONTROL THE RIPENING OF PAPAYA FRUIT AND
; TITLE OF INVENTION: CONFER DISEASE RESISTANCE TO PAPAYA PLANTS

```



; FILE REFERENCE: 19603/3601  
; CURRENT APPLICATION NUMBER: US/10/121,539  
; PRIOR FILING DATE: 2002-04-11  
; PRIOR FILING DATE: 2002-04-11  
; PRIOR FILING DATE: 2001-04-11  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 721  
; TYPE: PRT  
; ORGANISM: Carica papaya  
US-10-121-539-6

Query Match 54.5%; Score 48; DB 15; Length 721;  
Best Local Similarity 68.8%; Pred. No. 14;  
Matches 11; Conservative 1; Mismatches 0; Indels 4; Gaps 0;

Qy 4 LRAGFYAVSFSLVAVG 19  
Db 526 LRAGVNVKVSLSIAVG 541

RESULT 7  
US-10-369-493-10510  
; Sequence 10510, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 10510  
; LENGTH: 473  
; TYPE: PRT  
; ORGANISM: Spingomonas aromaticivorans  
US-10-369-493-10510

Query Match 50.6%; Score 44.5; DB 15; Length 473;  
Best Local Similarity 50.0%; Pred. No. 34;  
Matches 10; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

Qy 3 LLRAGF---YAVSFSLVAVG 19  
Db 95 IAKAGFETRFVAVSFIRAAIG 114

RESULT 8  
US-09-847-102A-42  
; Sequence 42, Application US/09847102A  
; Publication No. US2003004409A1  
; GENERAL INFORMATION:  
; APPLICANT: University of California  
; APPLICANT: Carson, Dennis A.  
; APPLICANT: Carr, Maripat  
; APPLICANT: Rhee, Chae-Seo  
; APPLICANT: Lorenzo, Leonil M.  
; APPLICANT: Malini, Sen  
; TITLE OF INVENTION: IMMUNOLOGIC COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: STUDYING AND TREATING CANCERS EXPRESSING FRIZZLED ANTIGENS  
; FILE REFERENCE: 22000-20629.00  
; CURRENT APPLICATION NUMBER: US/09/847,102A  
; CURRENT FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 42  
; LENGTH: 536  
; TYPE: PRT  
; ORGANISM: Drosophila  
US-09-847-102A-42

Query Match 48.9%; Score 43; DB 10; Length 536;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LLLRAGFYAVSFSLVAVG 19  
Db 424 LMLRIGFFSGLFILPAVG 441

RESULT 9  
US-09-738-626-4562  
; Sequence 4562, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4562  
; LENGTH: 578  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4562

Query Match 48.9%; Score 43; DB 9; Length 578;  
Best Local Similarity 52.6%; Pred. No. 74;  
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 VLLLRAGFYAVSFSLVAVG 19  
Db 265 VLILGTGYLVNIDLVTVG 283

RESULT 10  
US-09-740-041-2  
; Sequence 2, Application US/09740041  
; Patent No. US20020082190A1  
; GENERAL INFORMATION:  
; APPLICANT: MERKULOV, Karl et al  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL001001  
; CURRENT APPLICATION NUMBER: US/09/740,041  
; CURRENT FILING DATE: 2000-12-20  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 589

```
; TYPE: PRT
; ORGANISM: Human
US-09-740-041-2

Query Match          47.2%; Score 41.5; DB 9; Length 589;
Best Local Similarity 52.2%; Pred. No. 1.4e+02;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy  2 LLLRAGF-----YAVSFSLVAVG 19
    |||||
Db  403 LLLVVGFSHTKGVAISFLVAVG 425
    |||||

RESULT 11
US-10-389-967-2
; Sequence 2, Application US/10389967
; Publication No. US20030166153A1
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO01001-DIV
; CURRENT APPLICATION NUMBER: US/10/389,967
; CURRENT FILING DATE: 2003-03-18
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-389-967-2

Query Match          47.2%; Score 41.5; DB 14; Length 589;
Best Local Similarity 52.2%; Pred. No. 1.4e+02;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy  2 LLLRAGF-----YAVSFSLVAVG 19
    |||||
Db  403 LLLVVGFSHTKGVAISFLVAVG 425
    |||||

RESULT 12
US-09-915-181A-3
; Sequence 3, Application US/09915181A
; Patent No. US20020098473A1
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ROBERT
; APPLICANT: BELLOCCHIO, ELIZABETH
; APPLICANT: FREMEAUX, ROBERT
; APPLICANT: REIMER, RICHARD
; TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
; FILE REFERENCE: 305T-932610US
; CURRENT APPLICATION NUMBER: US/09/915,181A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/220,556
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-181A-3

Query Match          47.2%; Score 41.5; DB 9; Length 850;
Best Local Similarity 52.2%; Pred. No. 2e+02;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

Qy  2 LLLRAGF-----YAVSFSLVAVG 19
    |||||
Db  416 LLLVVGFSHTKGVAISFLVAVG 438
    |||||
```

```
RESULT 13
US-09-738-626-6265
; Sequence 6265, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6265
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6265
```

```
Query Match          46.6%; Score 41; DB 9; Length 297;
Best Local Similarity 50.0%; Pred. No. 77;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy  1 VLLLRAGFYAVSFSLVAV 18
    |||||
Db  249 VILLELAIFALSFLGSAV 266
    |||||
```

```
RESULT 14
US-10-156-761-9631
; Sequence 9631, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9631
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9631
```

```
Query Match          46.6%; Score 41; DB 14; Length 320;
Best Local Similarity 57.1%; Pred. No. 84;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

Qy 2 LLLRAGFYAVSFLS 15  
 Db 190 LNLQAGFYMIFLA 203

RESULT 15  
 US-10-080-170-441  
 ; Sequence 441, Application US/10080170  
 ; Publication No. US20030129601A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: COLE, S.T.  
 ; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR  
 ; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR  
 ; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES  
 ; FILE REFERENCE: 03495.0218  
 ; CURRENT APPLICATION NUMBER: US/10/080,170  
 ; CURRENT FILING DATE: 2002-06-10  
 ; PRIOR APPLICATION NUMBER: 60/270,123  
 ; PRIOR FILING DATE: 2001-02-22  
 ; NUMBER OF SEQ ID NOS: 652  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 441  
 ; LENGTH: 333  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium tuberculosis  
 US-10-080-170-441

Query Match 46.6%; Score 41; DB 14; Length 333;  
 Best Local Similarity 55.6%; Pred. No. 88;  
 Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LLLRAGFYAVSFLS 19  
 Db 46 VLLRQGEPAVSFLISSG 63

Search completed: February 26, 2004, 12:33:55  
 Job time : 8.60212 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 3.67905 Seconds  
(without alignments)  
266.616 Million cell updates/sec

Title: US-10-036-444-5

Perfect score: 88

Sequence: 1 VLLLRAGFYAVSFLSVAVG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	53.4	838	2	US-08-696-944-19
2	44	50.0	87	4	US-09-107-532A-7259
3	44	50.0	385	4	US-09-491-577-20
4	44	50.0	602	4	US-09-134-001C-4416
5	42	47.7	435	4	US-09-252-991A-17750
6	41.5	47.2	589	4	US-08-740-041-2
7	40	45.5	263	4	US-09-134-000C-5703
8	40	45.5	272	4	US-09-252-991A-24088
9	40	45.5	385	4	US-09-328-352-5641
10	40	45.5	394	4	US-08-543-681A-5779
11	39	44.3	69	3	US-08-783-974-3
12	39	44.3	69	4	US-09-546-306-3
13	39	44.3	726	4	US-09-252-991A-28519
14	39	44.3	730	2	US-08-696-944-2
15	39	44.3	740	4	US-09-489-039A-11157
16	39	44.3	1263	4	US-09-351-224E-11
17	39	44.3	1263	4	US-09-677-488A-11
18	39	44.3	1263	4	US-09-677-682B-11
19	38.5	43.8	198	4	US-09-328-352-8198
20	38.5	43.8	560	1	US-08-647-484-2
21	38.5	43.8	560	1	US-08-647-481-2
22	38.5	43.8	560	1	US-08-430-033A-2
23	38.5	43.8	560	1	US-08-805-118-4
24	38.5	43.8	560	4	US-09-391-958-4
25	38.5	43.8	560	5	PCT-US96-05792-2
26	38.5	43.8	578	4	US-09-740-041-4
27	38	43.2	236	3	US-09-121-979-4

28	38	43.2	236	3	US-09-332-319-4	Sequence 4, Appli
29	38	43.2	236	4	US-09-239-857-2	Sequence 2, Appli
30	38	43.2	493	4	US-03-543-681A-4617	Sequence 4617, Ap
31	38	43.2	555	4	US-09-543-681A-4582	Sequence 4582, Ap
32	38	43.2	575	4	US-09-489-039A-10911	Sequence 10911, A
33	38	43.2	739	1	US-07-618-946B-22	Sequence 22, Appli
34	38	43.2	814	1	US-07-618-946B-23	Sequence 23, Appli
35	37.5	42.6	760	4	US-09-252-991A-18711	Sequence 18711, A
36	37	42.0	186	4	US-09-134-000C-4177	Sequence 4177, Ap
37	37	42.0	238	4	US-09-134-000C-3467	Sequence 3467, Ap
38	37	42.0	256	4	US-09-489-039A-11447	Sequence 11447, A
39	37	42.0	263	4	US-09-489-039A-13979	Sequence 13979, A
40	37	42.0	298	2	US-08-838-543-5	Sequence 5, Appli
41	37	42.0	336	1	US-08-118-270-50	Sequence 50, Appli
42	37	42.0	336	5	PCT-US93-08528-50	Sequence 50, Appli
43	37	42.0	349	4	US-09-489-039A-11026	Sequence 11026, A
44	37	42.0	376	1	US-08-614-801A-6	Sequence 6, Appli
45	37	42.0	404	4	US-08-630-915A-24	Sequence 24, Appli

## ALIGNMENTS

RESULT 1  
US-08-696-944-19  
; Sequence 19, Application US/08696944  
; Patent No. 5981831  
; GENERAL INFORMATION:  
; APPLICANT: Sumant CHENGAPPA  
; APPLICANT: Susan A. HELLYER  
; APPLICANT: John S. REID  
; APPLICANT: Jacqueline DE SILVA  
; TITLE OF INVENTION: No. 5981831el Exo-(1-4)-Beta-D Galactanase  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MS Word  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/696,944  
; FILING DATE: 23-AUG-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/00372  
; FILING DATE: 23-FEB-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9403423.8  
; FILING DATE: 23-FEB-1994  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 838 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-696-944-19

Query Match 53.4%; Score 47; DB 2; Length 838;

Best Local Similarity 62.5%; Pred. No. 9;  
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFLSVAVG 19

Db 530 LRAGVKNKISLISIAVG 545

RESULT 2

US-09-107-532A-7259  
; Sequence 7259, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; City: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 7259:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 87 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...87  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7259:  
US-09-107-532A-7259  
Query Match 50.0%; Score 44; DB 4; Length 87;  
Best Local Similarity 60.0%; Pred. No. 2.2;  
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 VLLLRAGFYAVSFLS 15  
||:|||||::||  
DB 15 VLMIRAGFPATITLS 29  
RESULT 3  
US-09-491-577-20  
; Sequence 20, Application US/09491577  
; Patent No. 6610511  
; GENERAL INFORMATION:  
; APPLICANT: Yale University  
; APPLICANT: Carlson, John R.  
; APPLICANT: Kim, Hunhyong  
; APPLICANT: Clyne, Peter J.  
; APPLICANT: Warr, Coral G.  
; TITLE OF INVENTION: No. 6610511el Family of Odorant Receptor Genes in Drosophila  
; FILE REFERENCE: 44574-5061-US  
; CURRENT APPLICATION NUMBER: US/09/491,577  
; CURRENT FILING DATE: 2000-01-25

; EARLIER APPLICATION NUMBER: US 60/117,132  
; EARLIER FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 385  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-09-491-577-20  
Query Match 50.0%; Score 44; DB 4; Length 385;  
Best Local Similarity 60.0%; Pred. No. 12;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 VLLLRAGFYAVSFLS 15  
||:|||||::||  
DB 264 VLVLTANFYAIVLS 278  
RESULT 4  
US-09-134-001C-4416  
; Sequence 4416, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4416  
; LENGTH: 602  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4416  
Query Match 50.0%; Score 44; DB 4; Length 602;  
Best Local Similarity 66.7%; Pred. No. 20;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 7 GFYAVSFLSVAV 18  
||||:||||:  
DB 272 GFYAVSFLFLV 283  
RESULT 5  
US-09-252-991A-17750  
; Sequence 17750, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17750  
; LENGTH: 435  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17750  
Query Match 47.7%; Score 42; DB 4; Length 435;

Thu Feb 26 12:38:26 2004

```
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 RAGFYAVSFSLVAVG 19
DB 204 RASFFAVLAVLVG 218

RESULT 6
US-09-740-041-2
; Sequence 2, Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CLO01001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Human
US-09-740-041-2

Query Match 47.2%; Score 41.5; DB 4; Length 589;
Best Local Similarity 52.2%; Pred. No. 52;
Matches 12; Conservative 2; Mismatches 4; Indels 5; Gaps 1;

QY 2 LLLRAGF-----YAVSFSLVAVG 19
DB 403 LLLVGFSGHKGVAISFLVLAVG 425

RESULT 7
US-09-134-000C-5703
; Sequence 5703, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5703
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5703

Query Match 45.5%; Score 40; DB 4; Length 263;
Best Local Similarity 46.2%; Pred. No. 37;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 GFYAVSFSLVAVG 19
DB 223 GWYIVGFLAIG 235

RESULT 8
US-09-252-991A-24088
; Sequence 24088, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24088
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24088

Query Match 45.5%; Score 40; DB 4; Length 272;
Best Local Similarity 56.2%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 LRAGFYAVSFSLVAVG 19
DB 151 VVYSLQAVGFLSVLGV 166

RESULT 9
US-09-328-352-5641
; Sequence 5641, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5641
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5641

Query Match 45.5%; Score 40; DB 4; Length 385;
Best Local Similarity 43.8%; Pred. No. 58;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 VLLLRAGFYAVSFSLV 16
DB 14 LLLVSGAFYSIDFLKV 29

RESULT 10
US-09-543-681A-5779
; Sequence 5779, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5779
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5779

Query Match 45.5%; Score 40; DB 4; Length 394;
```



```
Best Local Similarity 69.2%; Pred. No. 1.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

QY 2 LLLRAGFYAVSFL 14
Db 274 LLLRGHLAGVFL 286

RESULT 14
US-08-696-944-2
; Sequence 2, Application US/08696944
; Patent No. 5981831
; GENERAL INFORMATION:
; APPLICANT: Sumant CHENGAPPA
; APPLICANT: Susan A. HELDYER
; APPLICANT: John S. REID
; APPLICANT: Jacqueline DE SILVA
; TITLE OF INVENTION: No. 5981831el Exo-(1-4)-Beta-D Galactanase
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/696,944
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/00372
; FILING DATE: 23-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9403423.8
; FILING DATE: 23-FEB-1994
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 730 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-696-944-2

Query Match 44.3%; Score 39; DB 2; Length 730;
Best Local Similarity 56.2%; Pred. No. 1.8e+02; Mismatches 5; Indels 0; Gaps 0;
Matches 9; Conservative 2;

QY 4 LRAGFYAVSFLSVAG 19
Db 536 LRGNKISLLSVAG 551

RESULT 15
US-09-489-039A-11157
; Sequence 11157, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11157
; LENGTH: 740
```

```
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11157

Query Match 44.3%; Score 39; DB 4; Length 740;
Best Local Similarity 60.0%; Pred. No. 1.8e+02; Mismatches 5; Indels 0; Gaps 0;
Matches 9; Conservative 1;

QY 2 LLLRAGFYAVSFLSV 16
Db 654 LLIRCGFKANSFAGV 668

Search completed: February 26, 2004, 12:16:17
Job time : 5.67905 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 4.98939 Seconds  
(without alignments)  
636.214 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVPEPRCP 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: Pirl.\*

2: Pirl.\*

3: Pirl.\*

4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	60	29.4	1801	1	WRTS
2	59.5	28.2	660	1	QBEB3
3	59	28.9	544	2	T36645
4	58.5	28.7	116	2	T46473
5	58	28.4	180	2	E84768
6	57	27.9	435	2	S40993
7	56.5	27.7	1321	2	J80352
8	55	27.0	1620	2	T27283
9	54.5	26.7	147	2	S37485
10	54.5	26.7	3078	2	T28432
11	54	26.5	1300	2	A36502
12	53.5	26.2	2767	1	UHHU
13	53	26.0	460	2	D97679
14	53	26.0	597	2	T16006
15	52.5	25.7	91	2	S37486
16	52.5	25.7	153	2	PN0564
17	52.5	25.7	577	2	B37057
18	52.5	25.7	799	2	A38308
19	52	25.5	137	2	A87586
20	52	25.5	1798	2	S33869
21	51.5	25.2	106	2	E95330
22	51.5	25.2	138	2	T48808
23	51.5	25.2	378	2	B59180
24	51.5	25.2	425	2	T48724
25	51.5	25.2	436	2	D88826
26	51.5	25.2	438	2	S28263
27	51.5	25.2	929	2	T51932
28	51.5	25.2	1166	1	S06142
29	51.5	25.2	1344	2	S47412

## ALIGNMENTS

### RESULT 1

#### WRTS

laminin beta-2 chain precursor - rat

N/Alternate names: laminin chain B3; S-laminin

C/Species: Rattus norvegicus (Norway rat)

C/Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 10-Dec-1999

C/Accession: S03539

R/Hunter, D.D.; Shah, V.; Merlie, J.P.; Sanes, J.R.

Nature 338, 229-234, 1989

A/Title: A laminin-like adhesive protein concentrated in the synaptic cleft of the neur

A/Reference number: S03539; MUID:89159410; PMID:2922051

A/Accession: S03539

A/Molecule type: mRNA

A/Residues: 1-1801 <HUN>

A/Cross-references: EMBL:X16563; NID:G57250; PIDN:CAA34561.1; PID:G57251

A/Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin

C/Function:

A/Description: interact with cells and with other basement membrane proteins to promote

C/Superfamily: laminin beta-1 chain; laminin-type EGF-like homology

C/Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular

F/1-35/Domain: signal sequence #status predicted <SIG>

F/36-1801/Product: laminin beta-2 chain #status predicted <MAT>

F/36-285/Domain: VI <DOM5>

F/286-555/Domain: V <DOM5>

F/286-347/Domain: laminin-type EGF-like homology <LE01>

F/350-410/Domain: laminin-type EGF-like homology <LE02>

F/413-470/Domain: laminin-type EGF-like homology <LE03>

F/473-522/Domain: laminin-type EGF-like homology <LE04>

F/525-555/Domain: laminin-type EGF-like homology #status atypical <LE05>

F/556-784/Domain: IV <DOM4>

F/786-831/Domain: laminin-type EGF-like homology <LE06>

F/788-1196/Domain: III <DOM3>

F/834-877/Domain: laminin-type EGF-like homology <LE07>

F/880-927/Domain: laminin-type EGF-like homology <LE08>

F/930-986/Domain: laminin-type EGF-like homology <LE09>

F/989-1038/Domain: laminin-type EGF-like homology <LE10>

F/1041-1095/Domain: laminin-type EGF-like homology <LE11>

F/1098-1143/Domain: laminin-type EGF-like homology <LE12>

F/1146-1190/Domain: laminin-type EGF-like homology <LE13>

F/1197-1412/Domain: II <DOM2>

F/1197-1412/Region: heptad repeats

F/1413-1445/Domain: alpha <ALP>

F/1446-1801/Region: heptad repeats

F/1446-1801/Domain: I <DOM1>

F/45-50/Disulfide bonds: #status predicted

F/251-371.1088.1252.1311.1351.1502/Binding site: carbohydrate (Asn) (covalent) #status

F/1193.1196.1800/Disulfide bonds: interchain #status predicted

Query Match 29.4%; Score 60; DB 1; Length 1801;

Best Local Similarity 29.8%; Pred. No. 21;

Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

QY 5 YQKCHCHM-----GTHCHSSD-GPRGVIPERPC 32  
 Db 1113 FTGQCHCHAGRGTRCQCLHWDGFLQCRACDCDPRG-IDKPCQ 1158

RESULT 2

QOE3  
 BHLF1 protein - human herpesvirus 4 (strain B95-8)  
 C/Species: human herpesvirus 4, Epstein-Barr virus  
 C/Date: 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 23-Aug-1997  
 C/Accession: A03742  
 R/Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.  
 Mol. Biol. Med. 1, 21-45, 1983  
 A/Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus  
 A/Reference number: A93065; MUID:85035713; PMID:6092825  
 A/Accession: A03742  
 A/Molecule type: DNA  
 A/Residues: 1-660 <BAN>  
 R/Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H  
 Nature 310, 207-211, 1984  
 A/Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.  
 A/Reference number: A03794; MUID:84270667; PMID:6087149  
 A/Contents: annotation; protein coding region  
 C/Comment: The sequence contains four perfect repeats (residues 149-273, 274-398, 399-52  
 C/Superfamily: human herpesvirus 4 BHLF1 protein

Query Match 29.2%; Score 59.5; DB 1; Length 660;  
 Best Local Similarity 57.1%; Pred. No. 10;  
 Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

QY 13 MGTCHSSDGRGVIPERPC 33  
 Db 1 MGTQCQARGPR-TTLPHP 20

RESULT 3

T36645  
 probable large integral membrane protein - Streptomyces coelicolor  
 C/Species: Streptomyces coelicolor  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C/Accession: T36645  
 R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, June 1999  
 A/Reference number: Z21610  
 A/Accession: T36645  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-544 <OLI>  
 A/Cross-references: EMBL:AL078610; PIDN:CAB44413.1; GSPDB:GN00070; SCOEDB:SCH35.14c  
 A/Experimental source: strain A3(2)  
 C/Genetics:  
 A/Gene: SCOEDB:SCH35.14c

Query Match 28.9%; Score 59; DB 2; Length 544;  
 Best Local Similarity 52.6%; Pred. No. 9.9;  
 Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 12 HMGTHCHSSDGRGVIPERPC 30  
 Db 434 HFGTHVLGDGPPAVVPH 452

RESULT 4

T46473  
 hypothetical protein DKFZp434O1230.1 - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C/Accession: T46473  
 R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, January 2000  
 A/Reference number: Z23034  
 A/Accession: T46473  
 A/Status: preliminary

Query Match 27.9%; Score 57; DB 2; Length 435;  
 Best Local Similarity 38.7%; Pred. No. 14;  
 Matches 12; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 2 TVTYQQKCHCHMGTHCHSSDGRGVIPERPC 32  
 Db 324 TCSLSGKC-CKLNHC-----PDGTPETSC 348

RESULT 7

JEO352

A/Molecule type: mRNA  
 A/Residues: 1-116 <AAA>  
 A/Cross-references: EMBL:AL137489  
 A/Experimental source: adult testis; clone DKFZp434O1230  
 C/Genetics:  
 A/Note: DKFZp434O1230.1

Query Match 28.7%; Score 58.5; DB 2; Length 116;  
 Best Local Similarity 43.5%; Pred. No. 2.9;  
 Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 10 HCHMGTHCHSSDGRGVIPERPC 32  
 Db 51 HCHGLYRCHGRPQREG---PRC 70

RESULT 5

E84768  
 hypothetical protein At2g35430 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C/Accession: E84768  
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L  
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
 Nature 402, 761-768, 1999  
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A/Reference number: A84420; MUID:20083487; PMID:10617197  
 A/Accession: E84768  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-180 <STO>  
 A/Cross-references: GB:AE002093; NID:G3608145; PIDN:AAC36178.1; GSPDB:GN00139  
 C/Genetics:  
 A/Gene: At2g35430  
 A/Map position: 2

Query Match 28.4%; Score 58; DB 2; Length 180;  
 Best Local Similarity 42.9%; Pred. No. 4.9;  
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 5 YQKCHCHMGTHCHSSDGRG 25  
 Db 152 WQTTGYCPFGSHCHFAHPSG 172

RESULT 6

S40993  
 hypothetical protein K04H4.3 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 03-May-1994 #sequence\_revision 02-Aug-1994 #text\_change 02-Aug-1994  
 C/Accession: S40993  
 R/Ainscough, R.  
 submitted to the EMBL Data Library, October 1993  
 A/Reference number: S40991  
 A/Accession: S40993  
 A/Molecule type: DNA  
 A/Residues: 1-435 <AIN>  
 A/Cross-references: EMBL:Z27078  
 C/Genetics:  
 A/Introns: 90/1; 305/3; 379/1; 425/1

Query Match 27.9%; Score 57; DB 2; Length 435;  
 Best Local Similarity 38.7%; Pred. No. 14;  
 Matches 12; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 2 TVTYQQKCHCHMGTHCHSSDGRGVIPERPC 32  
 Db 324 TCSLSGKC-CKLNHC-----PDGTPETSC 348

RESULT 7

JEO352

mucin MUC5B, tracheobronchial - human  
 C:Species: Homo sapiens (man)  
 C>Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 21-Jul-2000  
 C:Accession: J030352  
 R:Offner, G.D.; Nunes, D.P.; Keates, A.C.; Afzal, N.H.; Troxler, R.F.  
 Biochem. Biophys. Res. Commun. 251, 350-355, 1998  
 A>Title: The amino-terminal sequence of MUC5B contains conserved multifunctional D domain  
 A:Reference number: J030352; MUID:99009274; PMID:9790959  
 A:Accession: J030352  
 A:Molecule type: mRNA  
 A:Residues: 1-1321 <OFF>  
 A:Cross-references: GB:A086604; NID:g3789926; PIDN:AAC67545.1; PID:g3789927  
 C:Comment: This protein is large multimeric glycoproteins which is secreted by epithelial cells  
 C:Genetics:  
 A:Gene: MUC5B

Query Match 27.7%; Score 56.5; DB 2; Length 1321;  
 Best Local Similarity 39.0%; Pred. No. 44;  
 Matches 16; Conservative 0; Mismatches 10; Indels 15; Gaps 3;

QY 7 GKCH-----CHMGTHCHS-----SDGPRGVIPRCP 33  
 DB 816 GSCHTLDVGC-FSTHCVSVCVCPPLVSDGSGGCAEEDCP 855

RESULT 8  
 T27283  
 hypothetical protein Y64G10A.f - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T27283  
 R:Ainscough, R.  
 submitted to the EMBL Data Library, September 1999  
 A:Reference number: Z20336  
 A:Accession: T27283  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1620 <WIL>  
 A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CBSP:Y64G10A.f  
 A:Experimental source: clone Y64G10A  
 C:Genetics:  
 A:Gene: CBSP:Y64G10A.f  
 A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 777/1; 816/1; 842/1; 855/1; 882/1; 908/1; 935/1; 962/1; 989/1; 1016/1; 1043/1; 1070/1; 1097/1; 1124/1; 1151/1; 1178/1; 1205/1; 1232/1; 1259/1; 1286/1; 1313/1; 1340/1; 1367/1; 1394/1; 1421/1; 1448/1; 1475/1; 1502/1; 1529/1; 1556/1; 1583/1; 1610/1; 1637/1; 1664/1; 1691/1; 1718/1; 1745/1; 1772/1; 1799/1; 1826/1; 1853/1; 1880/1; 1907/1; 1934/1; 1961/1; 1988/1; 2015/1; 2042/1; 2069/1; 2096/1; 2123/1; 2150/1; 2177/1; 2204/1; 2231/1; 2258/1; 2285/1; 2312/1; 2339/1; 2366/1; 2393/1; 2420/1; 2447/1; 2474/1; 2501/1; 2528/1; 2555/1; 2582/1; 2609/1; 2636/1; 2663/1; 2690/1; 2717/1; 2744/1; 2771/1; 2798/1; 2825/1; 2852/1; 2879/1; 2906/1; 2933/1; 2960/1; 2987/1; 3014/1; 3041/1; 3068/1; 3095/1; 3122/1; 3149/1; 3176/1; 3203/1; 3230/1; 3257/1; 3284/1; 3311/1; 3338/1; 3365/1; 3392/1; 3419/1; 3446/1; 3473/1; 3500/1; 3527/1; 3554/1; 3581/1; 3608/1; 3635/1; 3662/1; 3689/1; 3716/1; 3743/1; 3770/1; 3797/1; 3824/1; 3851/1; 3878/1; 3905/1; 3932/1; 3959/1; 3986/1; 4013/1; 4040/1; 4067/1; 4094/1; 4121/1; 4148/1; 4175/1; 4202/1; 4229/1; 4256/1; 4283/1; 4310/1; 4337/1; 4364/1; 4391/1; 4418/1; 4445/1; 4472/1; 4499/1; 4526/1; 4553/1; 4580/1; 4607/1; 4634/1; 4661/1; 4688/1; 4715/1; 4742/1; 4769/1; 4796/1; 4823/1; 4850/1; 4877/1; 4904/1; 4931/1; 4958/1; 4985/1; 5012/1; 5039/1; 5066/1; 5093/1; 5120/1; 5147/1; 5174/1; 5201/1; 5228/1; 5255/1; 5282/1; 5309/1; 5336/1; 5363/1; 5390/1; 5417/1; 5444/1; 5471/1; 5498/1; 5525/1; 5552/1; 5579/1; 5606/1; 5633/1; 5660/1; 5687/1; 5714/1; 5741/1; 5768/1; 5795/1; 5822/1; 5849/1; 5876/1; 5903/1; 5930/1; 5957/1; 5984/1; 6011/1; 6038/1; 6065/1; 6092/1; 6119/1; 6146/1; 6173/1; 6200/1; 6227/1; 6254/1; 6281/1; 6308/1; 6335/1; 6362/1; 6389/1; 6416/1; 6443/1; 6470/1; 6497/1; 6524/1; 6551/1; 6578/1; 6605/1; 6632/1; 6659/1; 6686/1; 6713/1; 6740/1; 6767/1; 6794/1; 6821/1; 6848/1; 6875/1; 6902/1; 6929/1; 6956/1; 6983/1; 7010/1; 7037/1; 7064/1; 7091/1; 7118/1; 7145/1; 7172/1; 7199/1; 7226/1; 7253/1; 7280/1; 7307/1; 7334/1; 7361/1; 7388/1; 7415/1; 7442/1; 7469/1; 7496/1; 7523/1; 7550/1; 7577/1; 7604/1; 7631/1; 7658/1; 7685/1; 7712/1; 7739/1; 7766/1; 7793/1; 7820/1; 7847/1; 7874/1; 7901/1; 7928/1; 7955/1; 7982/1; 8009/1; 8036/1; 8063/1; 8090/1; 8117/1; 8144/1; 8171/1; 8198/1; 8225/1; 8252/1; 8279/1; 8306/1; 8333/1; 8360/1; 8387/1; 8414/1; 8441/1; 8468/1; 8495/1; 8522/1; 8549/1; 8576/1; 8603/1; 8630/1; 8657/1; 8684/1; 8711/1; 8738/1; 8765/1; 8792/1; 8819/1; 8846/1; 8873/1; 8900/1; 8927/1; 8954/1; 8981/1; 9008/1; 9035/1; 9062/1; 9089/1; 9116/1; 9143/1; 9170/1; 9197/1; 9224/1; 9251/1; 9278/1; 9305/1; 9332/1; 9359/1; 9386/1; 9413/1; 9440/1; 9467/1; 9494/1; 9521/1; 9548/1; 9575/1; 9602/1; 9629/1; 9656/1; 9683/1; 9710/1; 9737/1; 9764/1; 9791/1; 9818/1; 9845/1; 9872/1; 9899/1; 9926/1; 9953/1; 9980/1; 10007/1; 10034/1; 10061/1; 10088/1; 10115/1; 10142/1; 10169/1; 10196/1; 10223/1; 10250/1; 10277/1; 10304/1; 10331/1; 10358/1; 10385/1; 10412/1; 10439/1; 10466/1; 10493/1; 10520/1; 10547/1; 10574/1; 10601/1; 10628/1; 10655/1; 10682/1; 10709/1; 10736/1; 10763/1; 10790/1; 10817/1; 10844/1; 10871/1; 10898/1; 10925/1; 10952/1; 10979/1; 11006/1; 11033/1; 11060/1; 11087/1; 11114/1; 11141/1; 11168/1; 11195/1; 11222/1; 11249/1; 11276/1; 11303/1; 11330/1; 11357/1; 11384/1; 11411/1; 11438/1; 11465/1; 11492/1; 11519/1; 11546/1; 11573/1; 11600/1; 11627/1; 11654/1; 11681/1; 11708/1; 11735/1; 11762/1; 11789/1; 11816/1; 11843/1; 11870/1; 11897/1; 11924/1; 11951/1; 11978/1; 12005/1; 12032/1; 12059/1; 12086/1; 12113/1; 12140/1; 12167/1; 12194/1; 12221/1; 12248/1; 12275/1; 12302/1; 12329/1; 12356/1; 12383/1; 12410/1; 12437/1; 12464/1; 12491/1; 12518/1; 12545/1; 12572/1; 12599/1; 12626/1; 12653/1; 12680/1; 12707/1; 12734/1; 12761/1; 12788/1; 12815/1; 12842/1; 12869/1; 12896/1; 12923/1; 12950/1; 12977/1; 13004/1; 13031/1; 13058/1; 13085/1; 13112/1; 13139/1; 13166/1; 13193/1; 13220/1; 13247/1; 13274/1; 13301/1; 13328/1; 13355/1; 13382/1; 13409/1; 13436/1; 13463/1; 13490/1; 13517/1; 13544/1; 13571/1; 13598/1; 13625/1; 13652/1; 13679/1; 13706/1; 13733/1; 13760/1; 13787/1; 13814/1; 13841/1; 13868/1; 13895/1; 13922/1; 13949/1; 13976/1; 14003/1; 14030/1; 14057/1; 14084/1; 14111/1; 14138/1; 14165/1; 14192/1; 14219/1; 14246/1; 14273/1; 14300/1; 14327/1; 14354/1; 14381/1; 14408/1; 14435/1; 14462/1; 14489/1; 14516/1; 14543/1; 14570/1; 14597/1; 14624/1; 14651/1; 14678/1; 14705/1; 14732/1; 14759/1; 14786/1; 14813/1; 14840/1; 14867/1; 14894/1; 14921/1; 14948/1; 14975/1; 15002/1; 15029/1; 15056/1; 15083/1; 15110/1; 15137/1; 15164/1; 15191/1; 15218/1; 15245/1; 15272/1; 15299/1; 15326/1; 15353/1; 15380/1; 15407/1; 15434/1; 15461/1; 15488/1; 15515/1; 15542/1; 15569/1; 15596/1; 15623/1; 15650/1; 15677/1; 15704/1; 15731/1; 15758/1; 15785/1; 15812/1; 15839/1; 15866/1; 15893/1; 15920/1; 15947/1; 15974/1; 16001/1; 16028/1; 16055/1; 16082/1; 16109/1; 16136/1; 16163/1; 16190/1; 16217/1; 16244/1; 16271/1; 16298/1; 16325/1; 16352/1; 16379/1; 16406/1; 16433/1; 16460/1; 16487/1; 16514/1; 16541/1; 16568/1; 16595/1; 16622/1; 16649/1; 16676/1; 16703/1; 16730/1; 16757/1; 16784/1; 16811/1; 16838/1; 16865/1; 16892/1; 16919/1; 16946/1; 16973/1; 17000/1; 17027/1; 17054/1; 17081/1; 17108/1; 17135/1; 17162/1; 17189/1; 17216/1; 17243/1; 17270/1; 17297/1; 17324/1; 17351/1; 17378/1; 17405/1; 17432/1; 17459/1; 17486/1; 17513/1; 17540/1; 17567/1; 17594/1; 17621/1; 17648/1; 17675/1; 17702/1; 17729/1; 17756/1; 17783/1; 17810/1; 17837/1; 17864/1; 17891/1; 17918/1; 17945/1; 17972/1; 18000/1; 18027/1; 18054/1; 18081/1; 18108/1; 18135/1; 18162/1; 18189/1; 18216/1; 18243/1; 18270/1; 18297/1; 18324/1; 18351/1; 18378/1; 18405/1; 18432/1; 18459/1; 18486/1; 18513/1; 18540/1; 18567/1; 18594/1; 18621/1; 18648/1; 18675/1; 18702/1; 18729/1; 18756/1; 18783/1; 18810/1; 18837/1; 18864/1; 18891/1; 18918/1; 18945/1; 18972/1; 19000/1; 19027/1; 19054/1; 19081/1; 19108/1; 19135/1; 19162/1; 19189/1; 19216/1; 19243/1; 19270/1; 19297/1; 19324/1; 19351/1; 19378/1; 19405/1; 19432/1; 19459/1; 19486/1; 19513/1; 19540/1; 19567/1; 19594/1; 19621/1; 19648/1; 19675/1; 19702/1; 19729/1; 19756/1; 19783/1; 19810/1; 19837/1; 19864/1; 19891/1; 19918/1; 19945/1; 19972/1; 20000/1; 20027/1; 20054/1; 20081/1; 20108/1; 20135/1; 20162/1; 20189/1; 20216/1; 20243/1; 20270/1; 20297/1; 20324/1; 20351/1; 20378/1; 20405/1; 20432/1; 20459/1; 20486/1; 20513/1; 20540/1; 20567/1; 20594/1; 20621/1; 20648/1; 20675/1; 20702/1; 20729/1; 20756/1; 20783/1; 20810/1; 20837/1; 20864/1; 20891/1; 20918/1; 20945/1; 20972/1; 21000/1; 21027/1; 21054/1; 21081/1; 21108/1; 21135/1; 21162/1; 21189/1; 21216/1; 21243/1; 21270/1; 21297/1; 21324/1; 21351/1; 21378/1; 21405/1; 21432/1; 21459/1; 21486/1; 21513/1; 21540/1; 21567/1; 21594/1; 21621/1; 21648/1; 21675/1; 21702/1; 21729/1; 21756/1; 21783/1; 21810/1; 21837/1; 21864/1; 21891/1; 21918/1; 21945/1; 21972/1; 22000/1; 22027/1; 22054/1; 22081/1; 22108/1; 22135/1; 22162/1; 22189/1; 22216/1; 22243/1; 22270/1; 22297/1; 22324/1; 22351/1; 22378/1; 22405/1; 22432/1; 22459/1; 22486/1; 22513/1; 22540/1; 22567/1; 22594/1; 22621/1; 22648/1; 22675/1; 22702/1; 22729/1; 22756/1; 22783/1; 22810/1; 22837/1; 22864/1; 22891/1; 22918/1; 22945/1; 22972/1; 23000/1; 23027/1; 23054/1; 23081/1; 23108/1; 23135/1; 23162/1; 23189/1; 23216/1; 23243/1; 23270/1; 23297/1; 23324/1; 23351/1; 23378/1; 23405/1; 23432/1; 23459/1; 23486/1; 23513/1; 23540/1; 23567/1; 23594/1; 23621/1; 23648/1; 23675/1; 23702/1; 23729/1; 23756/1; 23783/1; 23810/1; 23837/1; 23864/1; 23891/1; 23918/1; 23945/1; 23972/1; 24000/1; 24027/1; 24054/1; 24081/1; 24108/1; 24135/1; 24162/1; 24189/1; 24216/1; 24243/1; 24270/1; 24297/1; 24324/1; 24351/1; 24378/1; 24405/1; 24432/1; 24459/1; 24486/1; 24513/1; 24540/1; 24567/1; 24594/1; 24621/1; 24648/1; 24675/1; 24702/1; 24729/1; 24756/1; 24783/1; 24810/1; 24837/1; 24864/1; 24891/1; 24918/1; 24945/1; 24972/1; 25000/1; 25027/1; 25054/1; 25081/1; 25108/1; 25135/1; 25162/1; 25189/1; 25216/1; 25243/1; 25270/1; 25297/1; 25324/1; 25351/1; 25378/1; 25405/1; 25432/1; 25459/1; 25486/1; 25513/1; 25540/1; 25567/1; 25594/1; 25621/1; 25648/1; 25675/1; 25702/1; 25729/1; 25756/1; 25783/1; 25810/1; 25837/1; 25864/1; 25891/1; 25918/1; 25945/1; 25972/1; 26000/1; 26027/1; 26054/1; 26081/1; 26108/1; 26135/1; 26162/1; 26189/1; 26216/1; 26243/1; 26270/1; 26297/1; 26324/1; 26351/1; 26378/1; 26405/1; 26432/1; 26459/1; 26486/1; 26513/1; 26540/1; 26567/1; 26594/1; 26621/1; 26648/1; 26675/1; 26702/1; 26729/1; 26756/1; 26783/1; 26810/1; 26837/1; 26864/1; 26891/1; 26918/1; 26945/1; 26972/1; 27000/1; 27027/1; 27054/1; 27081/1; 27108/1; 27135/1; 27162/1; 27189/1; 27216/1; 27243/1; 27270/1; 27297/1; 27324/1; 27351/1; 27378/1; 27405/1; 27432/1; 27459/1; 27486/1; 27513/1; 27540/1; 27567/1; 27594/1; 27621/1; 27648/1; 27675/1; 27702/1; 27729/1; 27756/1; 27783/1; 27810/1; 27837/1; 27864/1; 27891/1; 27918/1; 27945/1; 27972/1; 28000/1; 28027/1; 28054/1; 28081/1; 28108/1; 28135/1; 28162/1; 28189/1; 28216/1; 28243/1; 28270/1; 28297/1; 28324/1; 28351/1; 28378/1; 28405/1; 28432/1; 28459/1; 28486/1; 28513/1; 28540/1; 28567/1; 28594/1; 28621/1; 28648/1; 28675/1; 28702/1; 28729/1; 28756/1; 28783/1; 28810/1; 28837/1; 28864/1; 28891/1; 28918/1; 28945/1; 28972/1; 29000/1; 29027/1; 29054/1; 29081/1; 29108/1; 29135/1; 29162/1; 29189/1; 29216/1; 29243/1; 29270/1; 29297/1; 29324/1; 29351/1; 29378/1; 29405/1; 29432/1; 29459/1; 29486/1; 29513/1; 29540/1; 29567/1; 29594/1; 29621/1; 29648/1; 29675/1; 29702/1; 29729/1; 29756/1; 29783/1; 29810/1; 29837/1; 29864/1; 29891/1; 29918/1; 29945/1; 29972/1; 30000/1; 30027/1; 30054/1; 30081/1; 30108/1; 30135/1; 30162/1; 30189/1; 30216/1; 30243/1; 30270/1; 30297/1; 30324/1; 30351/1; 30378/1; 30405/1; 30432/1; 30459/1; 30486/1; 30513/1; 30540/1; 30567/1; 30594/1; 30621/1; 30648/1; 30675/1; 30702/1; 30729/1; 30756/1; 30783/1; 30810/1; 30837/1; 30864/1; 30891/1; 30918/1; 30945/1; 30972/1; 31000/1; 31027/1; 31054/1; 31081/1; 31108/1; 31135/1; 31162/1; 31189/1; 31216/1; 31243/1; 31270/1; 31297/1; 31324/1; 31351/1; 31378/1; 31405/1; 31432/1; 31459/1; 31486/1; 31513/1; 31540/1; 31567/1; 31594/1; 31621/1; 31648/1; 31675/1; 31702/1; 31729/1; 31756/1; 31783/1; 31810/1; 31837/1; 31864/1; 31891/1; 31918/1; 31945/1; 31972/1; 32000/1; 32027/1; 32054/1; 32081/1; 32108/1; 32135/1; 32162/1; 32189/1; 32216/1; 32243/1; 32270/1; 32297/1; 32324/1; 32351/1; 32378/1; 32405/1; 32432/1; 32459/1; 32486/1; 32513/1; 32540/1; 32567/1; 32594/1; 32621/1; 32648/1; 32675/1; 32702/1; 32729/1; 32756/1; 32783/1; 32810/1; 32837/1; 32864/1; 32891/1; 32918/1; 32945/1; 32972/1; 33000/1; 33027/1; 33054/1; 33081/1; 33108/1; 33135/1; 33162/1; 33189/1; 33216/1; 33243/1; 33270/1; 33297/1; 33324/1; 33351/1; 33378/1; 33405/1; 33432/1; 33459/1; 33486/1; 33513/1; 33540/1; 33567/1; 33594/1; 33621/1; 33648/1; 33675/1; 33702/1; 33729/1; 33756/1; 33783/1; 33810/1; 33837/1; 33864/1; 33891/1; 33918/1; 33945/1; 33972/1; 34000/1; 34027/1; 34054/1; 34081/1; 34108/1; 34135/1; 34162/1; 34189/1; 34216/1; 34243/1; 34270/1; 34297/1; 34324/1; 34351/1; 34378/1; 34405/1; 34432/1; 34459/1; 34486/1; 34513/1; 34540/1; 34567/1; 34594/1; 34621/1; 34648/1; 34675/1; 34702/1; 34729/1; 34756/1; 34783/1; 34810/1; 34837/1; 34864/1; 34891/1; 34918/1; 34945/1; 34972/1; 35000/1; 35027/1; 35054/1; 35081/1; 35108/1; 35135/1; 35162/1; 35189/1; 35216/1; 35243/1; 35270/1; 35297/1; 35324/1; 35351/1; 35378/1; 35405/1; 35432/1; 35459/1; 35486/1; 35513/1; 35540/1; 35567/1; 35594/1; 35621/1; 35648/1; 35675/1; 35702/1; 35729/1;

submitted to the EMBL Data Library, April 1988  
A;Description: Human mRNA for thyroglobulin.  
A;Reference number: A59110

A;Accession: A59110  
A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-2767 <MAL1>

A;Cross-references: GB:X05615; NID:G37173; PIDN:CAA29104.1; PID:G37174

A;Note: revision to S00014

R;Walthery, Y.; Lissitzky, S.

Eur. J. Biochem. 165, 491-498, 1987

A;Title: Primary structure of human thyroglobulin deduced from the sequence of its 8448-

A;Reference number: S00014; MUID:8746630; PMID:3595599

A;Accession: S00014

A;Molecule type: mRNA

A;Residues: 1-1041,'Y', 1043-1057,'T', 1059-2767 <MAL2>

A;Cross-references: GB:X05615; NID:G37173

A;Note: this sequence is revised in A59110

R;Walthery, Y.; Lissitzky, S.

Eur. J. Biochem. 147, 53-58, 1985

A;Title: Sequence of the 5'-end quarter of the human-thyroglobulin messenger ribonucleic

A;Reference number: A01532; MUID:85127024; PMID:3971976

A;Accession: A01532

A;Molecule type: mRNA

A;Residues: 1-730 <MA>

A;Cross-references: GB:X02154; NID:G37175; PIDN:CAA26089.1; PID:gl335349

A;Note: the translated sequence in GenBank entry HSTHYR5, release 11.0, (PIDN:CAA26089

n 1-Mat

R;Parna, J.; Christophe, D.; Pohl, V.; Vassart, G.

J. Mol. Biol. 136, 769-779, 1987

A;Title: Structural organization of the 5' region of the thyroglobulin gene. Evidence for

A;Reference number: S03422; MUID:88062712; PMID:3681978

A;Accession: S03422

A;Molecule type: DNA

A;Residues: 1-134,'Q', 136-415,'G', 415-652,'G', 654-733,'A', 735-737,'880-983','DR', 985-999 <PAR>

A;Cross-references: EMBL:X06059; NID:G37145; PIDN:CAA29454.1; PID:gl3359884; EMBL:X06067;

R;Christophe, D.; Cabrer, B.; Bacolla, A.; Targovnik, H.; Pohl, V.; Vassart, G.

Nucleic Acids Res. 13, 5127-5144, 1985

A;Title: An unusually long poly(purine)-poly(pyrimidine) sequence is located upstream of

A;Reference number: I38343; MUID:85269632; PMID:2991855

A;Accession: I38343

A;Status: translation not shown; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-22,'GRF', <CHR>

A;Cross-references: EMBL:X02749; NID:G37162; PIDN:CAA26527.1; PID:g758106

R;Targovnik, H.M.; Cochaux, P.; Corach, D.; Vassart, G.

Mol. Cell. Endocrinol. 84, R23-R26, 1992

A;Title: Identification of a minor 19 mRNA transcript in RNA from normal and goitrous th

A;Reference number: I57669; MUID:92347597; PMID:1639210

A;Accession: I57669

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1503-1508,'L', 1567-1601 <REW>

A;Cross-references: GB:S40607; NID:G252170; PIDN:AA822685.1; PID:G252171

A;Note: this sequence fragment represents a minor splice form

R;Marriq, C.; Lejeune, P.J.; Venot, N.; Vinet, L.

FEBS Lett. 242, 414-418, 1989

A;Title: Hormone synthesis in human thyroglobulin: possible cleavage of the polypeptide

A;Reference number: S02266; MUID:8912111; PMID:2914619

A;Contents: evidence for Tyr-149 as ring donor for thyroxine formation

A;Accession: S02266

A;Molecule type: protein

A;Residues: 101-105;114-121;126-131;143-149 <WAR>

A;Note: only the first peptide was sequenced; others were isolated and their amino acid

R;Gentile, F.; Salvatore, G.

Eur. J. Biochem. 218, 603-621, 1993

A;Title: Preferential sites of proteolytic cleavage of bovine, human and rat thyroglobul

A;Reference number: S39431; MUID:94094855; PMID:8269951

A;Accession: S39431

A;Molecule type: Protein

A;Residues: 20-27;522-527;541-547;616-619,'X', 621;1001-1005;1009-1011,'X', 1013;1424-1430

R;Xiao, S.; Pollock, H.G.; Taurag, A.; Rawitch, A.B.

Arch. Biochem. Biophys. 320, 96-105, 1995

A;Title: Characterization of hormonogenic sites in an N-terminal, cyanogen bromide frag  
A;Reference number: S66241; MUID:95314327; PMID:7793989

A;Accession: S66241

A;Molecule type: protein

A;Residues: 'D', 21-23;45-47;147-148;150-152 <XIA>

R;Yang, S.X.; Pollock, H.G.; Rawitch, A.B.

Arch. Biochem. Biophys. 327, 61-70, 1996

A;Title: Glycosylation in human thyroglobulin: location of the N-linked oligosaccharide

A;Reference number: S62778; MUID:96201348; PMID:8615697

A;Accession: S62778

A;Molecule type: protein

A;Residues: 65-84;136-210,'T', 212-213;476-492;523-539;741-770;811-848;850-853;938-950;1

6-1783,'D', 1785,'R', 1999-2017;2241-2258;2270-2285,'P', 2287-2307;2578-2584 <YAN>

R;Ieliri, T.; Cochaux, P.; Targovnik, H.M.; Suzuki, M.; Shimoda, S.; Perret, J.; Vassart

J. Clin. Invest. 88, 1901-1905, 1991

A;Title: A 3' splice site mutation in the thyroglobulin gene responsible for congenital

A;Reference number: I55565; MUID:92091498; PMID:1752952

A;Accession: I55565

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 90-91,160-161 <IEI>

A;Cross-references: GB:S71821; NID:G240935; PIDN:AA820665.1; PID:G240936

A;Note: mutant splice form

C;Comment: The thyroglobulin molecule is produced in the thyroid gland and is the precu

C;Genetics:

A;Gene: GDB:TG

A;Cross-references: GDB:I20434; OMIM:188450

A;Map position: 8q24.2-8q24.3

A;Introns: 23/1; 59/2; 92/1; 160/1; 213/2; 249/1; 297/1; 359/1; 726/1; 921/1; 1000/1 #8

C;Complex: homodimer

C;Function:

A;Description: precursor of thyroid hormones thyroxine and triiodothyronine; iodine stc

C;Superfamily: thyroglobulin; cholinesterase homology; glycoprotein; homodimer; iodine; thyroid

C;Keywords: alternative splicing; duplication; thyroglobulin type I repeat homc

F;1-19/Dmain: signal sequence #status predicted <SIG>

F;20-2767/Product: thyroglobulin #status predicted <MAT>

F;34-92/Dmain: thyroglobulin type I repeat homology <THV1>

F;96-160/Dmain: thyroglobulin type I repeat homology <THY2>

F;164-297/Dmain: thyroglobulin type I repeat homology #status atypical <THV3>

F;301-358/Dmain: thyroglobulin type I repeat homology <THY4>

F;608-658/Dmain: thyroglobulin type I repeat homology <THY5>

F;662-726/Dmain: thyroglobulin type I repeat homology <THY6>

F;730-921/Dmain: thyroglobulin type I repeat homology #status atypical <THY7>

F;925-1072/Dmain: thyroglobulin type I repeat homology #status atypical <THY8>

F;1076-1144/Dmain: thyroglobulin type I repeat homology <THY9>

F;1148-1209/Dmain: thyroglobulin type I repeat homology <TH10>

F;1455-1468/Region: type II repeat

F;1469-1485/Region: type II repeat

F;1486-1502/Region: type II repeat

F;1602-1722/Region: type IIIa repeat

F;1723-1891/Region: type IIIb repeat

F;1892-1994/Region: type IIIa repeat

F;1995-2128/Region: type IIIb repeat

F;2129-2186/Region: type IIIa repeat

F;2227-2725/Dmain: cholinesterase homology <ACE>

F;24,2572,2586/Modified site: thyroxine (Tyr) #status predicted

F;76,198,484,529,748,816,947,1219,1348,1364,1715,1773,2012,2249,2294,2581/Binding site:

F;110,496,1868,2121/Binding site: carbohydrate (Asn) (covalent) #status absent

F;149/Modified site: dehydralanine (Tyr) #status predicted

F;2765/Modified site: triiodothyronine (Tyr) #status predicted

Query Match 26.2%; Score 53.5; DB 1; Length 2767;

Best Local Similarity 31.2%; Pred. No. 1.9e-02;

Matches 15; Conservative 2; Mismatches 12; Indels 19; Gaps 3;

QY 5 YQKGC-----HCHMGT-HCHSSDGRGVPE-----PRCP 33

DB 1026 YMPQCDAFSGSWPVCAGTGHWCVDKGGFPGSLTARSLLQIPQCP 1073

RESULT 13

D97679

argininosuccinate synthase (PA3525) [imported] - Agrobacterium tumefaciens (strain C58,

C:Species: Agrobacterium tumefaciens  
 C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C:Accession: D97679  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman, A.; Liu, F.; Wollan, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: D97679  
 A>Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-460 <KUR>  
 A:Cross-references: GB:AE007869; PIDN:AAK88389.1; PID:GI5157878; GSPDB:GN00169  
 C:Genetics:  
 A:Gene: AGR C 4836  
 A:Map position: circular chromosome  
 C:Superfamily: argininosuccinate synthase

Query Match 26.0%; Score 53; DB 2; Length 460;  
 Best Local Similarity 40.0%; Pred. No. 46;  
 Matches 12; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 4 YQKGKCHCMGTHCHSSDGRGVPEPCP 33  
 DB 18 YLTKRFTFGLHCKSAEIPR----KWRCP 43

## RESULT 14

T16006  
 protein kinase C3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jan-2000  
 C:Accession: T16006; T37253  
 R:Chisoe, S.  
 submitted to the EMBL Data Library, September 1995  
 A:Description: The sequence of C. elegans cosmid F09E5.  
 A:Reference number: Z18444  
 A:Accession: T16006  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-597 <CHI>  
 A:Cross-references: EMBL:U37429; NID:G1019949; PID:G1019950; PIDN:AAA79341.1; CESP:F09E5  
 A:Experimental source: strain Bristol N2  
 R:Wu, S.L.; Rubin, C.S.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: Molecular characterization of an atypical protein kinase C from C. elegans  
 A:Reference number: Z21647  
 A:Accession: T37253  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-597 <WUS>  
 A:Cross-references: EMBL:AF025666; PIDN:AAB88885.1  
 A:Experimental source: strain Bristol N2  
 C:Genetics:  
 A:Gene: F09E5.1; pkc-3  
 A:Map position: 11  
 A:Introns: 22/3; 109/1; 137/2; 179/3; 322/3; 442/2; 556/3  
 C:Superfamily: protein kinase C zeta; protein kinase C zinc-binding repeat homology; protein kinase C zinc-binding repeat homology <KZN>  
 F:251-522/Domain: protein kinase homology <KIN>

Query Match 26.0%; Score 53; DB 2; Length 597;  
 Best Local Similarity 50.0%; Pred. No. 58;  
 Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 2;

QY 8 KCHCMGTHC-HSSDGRGVIP 28  
 DB 168 KCHRVTHCGGALQGP-NIIP 188

## RESULT 15

S37486  
 gene msg3 protein - mouse

C:Species: Mus musculus (house mouse)  
 C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 11-Jan-2000  
 C:Accession: I48671; S37486  
 R:Tronik-Le Roux, D.; Senorale-Pose, M.; Rougeon, P.  
 Gene 142, 175-182, 1994  
 A:Title: Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice  
 A:Reference number: I48669; MUID:94252564; PMID:8194749  
 A:Accession: I48671  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-91 <RES>  
 A:Cross-references: EMBL:X71631; NID:G406260; PIDN:CRA50638.1; PID:G406261  
 C:Genetics:  
 A:Gene: msg3  
 C:Superfamily: proline-rich peptide P-B

Query Match 25.7%; Score 52.5; DB 2; Length 91;  
 Best Local Similarity 45.5%; Pred. No. 13;  
 Matches 10; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

QY 9 CHCHMGTHCHSSDGRGVPEP 30  
 DB 20 CECHRGPRRD---PRGPPFP 38

Search completed: February 26, 2004, 12:14:56  
 Job time : 7.98939 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 2.80106 Seconds  
(without alignments)  
613.452 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVIPEPRCP 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	29.4	1799	1 LMB2 MOUSE	Q61292 mus musculus
2	60	29.4	1801	1 LMB2 RAT	P15800 rattus norv
3	59.5	29.2	660	1 YHL1 EBV	P03181 epstein-bar
4	58.5	28.7	2471	1 NTC2 HUMAN	Q04721 homo sapien
5	54.5	26.7	147	1 SMR1 MOUSE	Q61900 mus musculus
6	54	26.5	1300	1 TRR_CAVPO	P14617 cavia porce
7	53.5	26.2	2768	1 THYG_HUMAN	P01266 homo sapien
8	52.5	25.7	577	1 ITB6_CAVPO	P18563 cavia porce
9	52.5	25.7	589	1 ZN44 HUMAN	P15621 homo sapien
10	52.5	25.7	799	1 ITB5 HUMAN	P18084 homo sapien
11	52.5	25.7	1565	1 DMN_HUMAN	O15061 homo sapien
12	52.5	25.7	2813	1 VWF_CANFA	Q28295 canis famil
13	52	25.5	1798	1 LMB2_HUMAN	P55258 homo sapien
14	51.5	25.2	147	1 HMR1_RHIME	P58278 rhizobium m
15	51.5	25.2	378	1 WIF1_BRARE	Q9W619 brachydanio
16	51.5	25.2	438	1 LIN3_CABEL	Q03345 caenorhabdi
17	51.5	25.2	461	1 Y514 HUMAN	O60269 homo sapien
18	51.5	25.2	1167	1 XMRK_XIPMA	P13388 xiphophorus
19	51.5	25.2	5703	1 MUSB_HUMAN	Q9HC84 homo sapien
20	51	25.0	178	1 NEL1_MOUSE	O61477 mus musculus
21	51	25.0	325	1 VT2_SFVKA	P25943 Shope fibro
22	51	25.0	398	1 DHZ2_XENLA	Q91611 xenopus lae
23	51	25.0	434	1 UROK_CHICK	P15120 gallus gall
24	51	25.0	699	1 ECM2_HUMAN	Q94769 homo sapien
25	51	25.0	2482	1 VNF_PIG	Q28833 sus scrofa
26	50.5	24.8	74	1 BDFE_HUMAN	O60594 homo sapien
27	50.5	24.8	106	1 Y402_METJA	Q57845 methanococc
28	50.5	24.8	515	1 KZAL_CABEL	Q3Xtd7 caenorhabdi
29	50.5	24.8	655	1 ITB5_PAPCI	Q07441 papio cynoc
30	50	24.5	396	1 DHH1_XENLA	Q91610 xenopus lae
31	50	24.5	759	1 MTS1_MOUSE	Q81814 mus musculus
32	50	24.5	5376	1 ZAN_MOUSE	Q88799 mus musculus
33	49	24.0	176	1 TRZ3_MOUSE	Q9er63 mus musculus

Q9h293 homo sapien  
P77365 escherichia  
P29825 myxoma viru  
O63371 rattus norv  
O14513 homo sapien  
P28737 cavia porce  
Q9qy70 z adam 15 p  
P77325 homo sapien  
O43052 schizosacch  
P58512 homo sapien  
P25117 sus scrofa  
P48467 neurospora

117E HUMAN  
YAFY ECOLI  
VT2 MYXVL  
P2Y6 RAT  
NAP5 HUMAN  
GRN\_CAVPO  
AD15 RAT  
AD08 HUMAN  
RGAL SCHPO  
CU67 HUMAN  
CALR\_PIG  
KINH\_NEUCR

49 24.0 177 1  
49 24.0 285 1  
49 24.0 326 1  
49 24.0 328 1  
49 24.0 416 1  
49 24.0 591 1  
49 24.0 816 1  
49 24.0 824 1  
49 24.0 1150 1  
48.5 23.8 204 1  
48.5 23.8 498 1  
48.5 23.8 928 1

ALIGNMENTS

RESULT 1

LMB2\_MOUSE

ID LMB2\_MOUSE STANDARD; PRT; 1799 AA.

AC Q61292; Q62182;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Laminin beta-2 chain precursor (S-laminin) (S-LAM).

CN LAMB2 OR LAMS

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=129/J;

RX MEDLINE=96278760; PubMed=8662701;

RA Durkin M.E., Gautam M., Loechel S., Sanes J.R., Merlie J.P.,

RA Albrechtsen R., Wever U.M.;

RT "Structural organization of the human and mouse laminin beta2 chain

RT genes, and alternative splicing at the 5' end of the human

RT transcript.";

RL J. Biol. Chem. 271:13407-13416(1996).

RN [2]

RP SEQUENCE OF 348-428 FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=94319092; PubMed=8043959;

RA Aberdam D., Galliano M.F., Mattei M.-G., Ortonne J.P., Meneguzzi G.;

RT "S-laminin gene (Lams) maps to F1 band of mouse chromosome 9.";

RL Mamm. Genome 5:393-394(1994).

RN [3]

RP FUNCTION.

RC STRAIN=129/J;

RX MEDLINE=95191650; PubMed=7885444;

RA Noakes P.G., Gautam M., Mudd J., Sanes J.R., Merlie J.P.;

RT "Aberrant differentiation of neuromuscular junctions in mice lacking

RT s-laminin/laminin beta 2.";

RL Nature 374:258-262(1995).

-!- FUNCTION: Binding to cells via a high affinity receptor, laminin

of cells into tissues during embryonic development By interacting

with other extracellular matrix components.

-!- FUNCTION: Laminin-3 (S-laminin) regulates the formation of motor

nerve terminals.

-!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three

different polypeptide chains (alpha, beta, gamma), which are bound

to each other by disulfide bonds into a cross-shaped molecule

comprising one long and three short arms with globules at each

end. The beta-2 chain is a subunit of laminin-3 (S-laminin),

laminin-4 (S-merosin), and laminin-7 (KS-laminin).

-!- SUBCELLULAR LOCATION: Extracellular; found in the basement

membranes (major component).

-!- TISSUE SPECIFICITY: Neuromuscular synapse and kidney glomerulus.

-!- DOMAIN: The alpha-helical domains I and II are thought to interact

with other laminin chains to form a coiled coil structure.

-!- DOMAIN: Domains VI and IV are globular.

DT	DISULFID	832	844	BY SIMILARITY.
DT	DISULFID	834	851	BY SIMILARITY.
DT	DISULFID	853	862	BY SIMILARITY.
DT	DISULFID	865	875	BY SIMILARITY.
DT	DISULFID	878	887	BY SIMILARITY.
DT	DISULFID	880	894	BY SIMILARITY.
DT	DISULFID	897	906	BY SIMILARITY.
DT	DISULFID	909	925	BY SIMILARITY.
DT	DISULFID	928	944	BY SIMILARITY.
DT	DISULFID	930	955	BY SIMILARITY.
DT	DISULFID	957	966	BY SIMILARITY.
DT	DISULFID	969	984	BY SIMILARITY.
DT	DISULFID	987	1001	BY SIMILARITY.
DT	DISULFID	989	1008	BY SIMILARITY.
DT	DISULFID	1023	1036	BY SIMILARITY.
DT	DISULFID	1096	1108	BY SIMILARITY.
DT	DISULFID	1098	1115	BY SIMILARITY.
DT	DISULFID	1117	1126	BY SIMILARITY.
DT	DISULFID	1129	1141	BY SIMILARITY.
DT	DISULFID	1144	1156	BY SIMILARITY.
DT	DISULFID	1146	1163	BY SIMILARITY.
DT	DISULFID	1165	1174	BY SIMILARITY.
DT	DISULFID	1177	1188	BY SIMILARITY.
DT	DISULFID	1191	1191	INTERCHAIN (PROBABLE).
DT	DISULFID	1194	1194	INTERCHAIN (PROBABLE).
DT	DISULFID	1798	1798	INTERCHAIN (PROBABLE).
DT	CARBOHYD	251	251	N-LINKED (GLCNAC. .)
DT	CARBOHYD	371	371	N-LINKED (GLCNAC. .)
DT	CARBOHYD	1086	1086	N-LINKED (GLCNAC. .)
DT	CARBOHYD	1250	1250	N-LINKED (GLCNAC. .)
DT	CARBOHYD	1309	1309	N-LINKED (GLCNAC. .)
DT	CARBOHYD	1349	1349	N-LINKED (GLCNAC. .)
DT	CARBOHYD	1500	1500	N-LINKED (GLCNAC. .)
DT	CARBOHYD	1799	1799	AA; 196352 MW; 1F28967A67AEDE33 CRC64;
DT	SEQUENCE			
DT	Query Match	29.4%;	Score 60;	DB 1; Length 1799;
DT	Best Local Similarity	29.8%;	Pred. No. 7.6;	
DT	Matches 14;	Conservative	5;	Mismatches 8; Indels 20; Gaps
DT	OY	5 YQGHCHQM-----	-----GTHCHSSD-GPRGVIPERC 32	
DT	DB	1111 FTGQCHAGFGGTCSECQELWGPGLQCRACDCDPRG-IDKPOC 1156		
DT	RESULT 2			
DT	ID_LMB2_RAT	STANDARD;	PRT;	1801 AA.
DT	AC	P15600;		
DT	DT	01-APR-1990 (Rel. 14, Created)		
DT	DT	01-APR-1990 (Rel. 14, Last sequence update)		
DT	DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DT	DT	Laminin beta-2 chain precursor (S-laminin) (Laminin chain B3).		
DT	GN	LAMB2.		
DT	GN	Rattus norvegicus (Rat).		
DT	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
DT	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
DT	OX	NCBI_TaxId=10116;		
DT	RP	[1]		
DT	RP	SEQUENCE FROM N.A.		
DT	RC	TISSUE=Liver;		
DT	RC	MEDLINE=89159410; PubMed=3922051;		
DT	RA	Hunter D.D., Shah V., Merlie J.P., Sanes J.R.;		
DT	RA	"A laminin-like adhesive protein concentrated in the synaptic cleft		
DT	RL	of the neuromuscular junction."		
DT	RL	Nature 338:229-234(1989).		
DT	CC	-!- FUNCTION: Binding to cells via a high affinity receptor, laminin		
DT	CC	is thought to mediate the attachment, migration and organization		
DT	CC	of cells into tissues during embryonic development by interacting		
DT	CC	with other extracellular matrix components.		
DT	CC	-!- SUBUNIT: Laminin is a complex glycoprotein, consisting of three		
DT	CC	different polypeptide chains (alpha, beta, gamma), which are bound		
DT	CC	to each other by disulfide bonds into a cross-shape molecule		



comprising one long and three short arms with globules at each end. The beta-2 chain is a subunit of laminin-3 (S-laminin), laminin-4 (S-merosin), and laminin-7 (KS-laminin).  
-!- SUBCELLULAR LOCATION: Extracellular.  
-!- TISSUE SPECIFICITY: Found in the basement membranes (major component). S-laminin is concentrated in the synaptic cleft of the neuromuscular junction.  
-!- DOMAIN: The alpha-helical domains I and II are thought to interact with other laminin chains to form a coiled coil structure.  
-!- DOMAIN: Domains VI and IV are globular.  
-!- SIMILARITY: Contains 1 laminin N-terminal domain.  
-!- SIMILARITY: Contains 13 laminin EGF-like domains.  
-!- SIMILARITY: Contains 1 laminin IV domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X16563; CAA34561.1; -.  
PIR; S03539; MURTS.  
HSP; P02468; 1KLO.  
InterPro; IPR006209; EGF like.  
InterPro; IPR002049; Laminin\_EGF.  
InterPro; IPR008211; LamNT.  
Pfam; PF00053; laminin\_EGF; 13.  
Pfam; PF00055; laminin\_Nterm; 1.  
PRINTS; PRO0011; EGF\_LAMININ.  
SMART; SM00180; EGF\_Lam; 13.  
SMART; SM00136; LamNT; 1.  
PROSITE; PS00022; EGF 1; 10.  
PROSITE; PS01186; EGF 2; 2.  
PROSITE; PS01248; LAMININ TYPE EGF; 12.  
Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
FT SIGNAL 1 35  
FT CHAIN 36 1801 LAMININ BETA-2 CHAIN.  
FT DOMAIN 36 283 LAMININ N-TERMINAL (DOMAIN VI).  
FT DOMAIN 286 349 LAMININ EGF-LIKE 1.  
FT DOMAIN 350 412 LAMININ EGF-LIKE 2.  
FT DOMAIN 413 472 LAMININ EGF-LIKE 3.  
FT DOMAIN 473 524 LAMININ EGF-LIKE 4.  
FT DOMAIN 525 555 LAMININ EGF-LIKE 5 (INCOMPLETE).  
FT DOMAIN 556 785 LAMININ DOMAIN IV.  
FT DOMAIN 786 833 LAMININ EGF-LIKE 6.  
FT DOMAIN 834 879 LAMININ EGF-LIKE 7.  
FT DOMAIN 880 929 LAMININ EGF-LIKE 8.  
FT DOMAIN 930 988 LAMININ EGF-LIKE 9.  
FT DOMAIN 989 1040 LAMININ EGF-LIKE 10.  
FT DOMAIN 1041 1097 LAMININ EGF-LIKE 11.  
FT DOMAIN 1098 1145 LAMININ EGF-LIKE 12.  
FT DOMAIN 1146 1192 LAMININ EGF-LIKE 13.  
FT DOMAIN 1193 1412 DOMAIN II.  
FT DOMAIN 1413 1445 DOMAIN ALPHA.  
FT DOMAIN 1446 1801 DOMAIN I.  
FT DOMAIN 1259 1306 COILED COIL (POTENTIAL).  
FT DOMAIN 1475 1529 COILED COIL (POTENTIAL).  
FT DOMAIN 1576 1793 COILED COIL (POTENTIAL).  
FT DISULFID 286 293 BY SIMILARITY.  
FT DISULFID 288 313 BY SIMILARITY.  
FT DISULFID 315 324 BY SIMILARITY.  
FT DISULFID 327 347 BY SIMILARITY.  
FT DISULFID 350 359 BY SIMILARITY.  
FT DISULFID 352 377 BY SIMILARITY.  
FT DISULFID 380 389 BY SIMILARITY.  
FT DISULFID 392 410 BY SIMILARITY.  
FT DISULFID 413 426 BY SIMILARITY.  
FT DISULFID 415 441 BY SIMILARITY.  
FT DISULFID 443 452 BY SIMILARITY.  
FT DISULFID 445 470 BY SIMILARITY.

487 473 DISULFID 473 486 BY SIMILARITY.  
494 475 DISULFID 475 493 BY SIMILARITY.  
505 496 DISULFID 496 504 BY SIMILARITY.  
522 508 DISULFID 508 520 BY SIMILARITY.  
798 786 DISULFID 786 796 BY SIMILARITY.  
805 788 DISULFID 788 803 BY SIMILARITY.  
816 807 DISULFID 807 816 BY SIMILARITY.  
831 819 DISULFID 819 831 BY SIMILARITY.  
846 834 DISULFID 834 846 BY SIMILARITY.  
853 846 DISULFID 846 853 BY SIMILARITY.  
854 853 DISULFID 853 854 BY SIMILARITY.  
877 867 DISULFID 867 877 BY SIMILARITY.  
889 880 DISULFID 880 889 BY SIMILARITY.  
896 882 DISULFID 882 896 BY SIMILARITY.  
908 899 DISULFID 899 908 BY SIMILARITY.  
927 911 DISULFID 911 927 BY SIMILARITY.  
946 930 DISULFID 930 946 BY SIMILARITY.  
957 932 DISULFID 932 957 BY SIMILARITY.  
968 959 DISULFID 959 968 BY SIMILARITY.  
986 971 DISULFID 971 986 BY SIMILARITY.  
1003 989 DISULFID 989 1003 BY SIMILARITY.  
1010 991 DISULFID 991 1010 BY SIMILARITY.  
1028 1013 DISULFID 1013 1028 BY SIMILARITY.  
1038 1025 DISULFID 1025 1038 BY SIMILARITY.  
1110 1098 DISULFID 1098 1110 BY SIMILARITY.  
1117 1100 DISULFID 1100 1117 BY SIMILARITY.  
1128 1119 DISULFID 1119 1128 BY SIMILARITY.  
1143 1131 DISULFID 1131 1143 BY SIMILARITY.  
1158 1146 DISULFID 1146 1158 BY SIMILARITY.  
1165 1148 DISULFID 1148 1165 BY SIMILARITY.  
1176 1167 DISULFID 1167 1176 BY SIMILARITY.  
1190 1179 DISULFID 1179 1190 BY SIMILARITY.  
1193 1193 DISULFID 1193 1193 INTERCHAIN (PROBABLE).  
1196 1196 DISULFID 1196 1196 INTERCHAIN (PROBABLE).  
1800 1800 DISULFID 1800 1800 INTERCHAIN (PROBABLE).  
251 251 DISULFID 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).  
371 371 DISULFID 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).  
1088 1088 DISULFID 1088 1088 N-LINKED (GLCNAC. . .) (POTENTIAL).  
1252 1252 DISULFID 1252 1252 N-LINKED (GLCNAC. . .) (POTENTIAL).  
1311 1311 DISULFID 1311 1311 N-LINKED (GLCNAC. . .) (POTENTIAL).  
1351 1351 DISULFID 1351 1351 N-LINKED (GLCNAC. . .) (POTENTIAL).  
1502 1502 DISULFID 1502 1502 N-LINKED (GLCNAC. . .) (POTENTIAL).  
1801 AA; 196473 MW; 97AEF32FBF31FA75 CRC64;  
Query Match 29.4%; Score 60; DB 1; Length 1801;  
Best Local Similarity 29.8%; Pred. No. 7.6;  
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;  
Qy 5 YQKCHCHM-----GTHCHSD-GPRGVPEPRC 32  
Db 1113 FTGCHCHAGFGGRTCTSECQELHWGDFGLQCRACDCDPRG-IDKPCQ 1158  
-----  
RESULT 3  
YHL1\_EBV  
ID YHL1\_EBV STANDARD; PRT; 660 AA.  
AC P03181;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Hypothetical BHLF1 protein.  
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Gammaherpesvirinae; Lymphocryptovirus.  
OX NCBI\_TaxID=10377;  
RN [1]  
RX SEQUENCE FROM N.A.  
RX MEDLINE=84270667; PubMed=6087149;  
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,  
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,  
RA Tuffnell P.S., Barrell B.G.;  
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";  
RT Nature 310:207-211(1984).



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V01555; -; NOT_ANNOTATED_CDS.
CC DR PIR; A03742; Q0B3.
CC KW Hypothetical protein; Early protein; Repeat.
CC FT DOMAIN 149 648 4 X 125 AA TANDEN REPEATS.
CC FT REPEAT 149 273 1.
CC FT REPEAT 274 398 2.
CC FT REPEAT 399 523 3.
CC FT REPEAT 524 648 4.
CC SQ SEQUENCE 660 AA; 66244 MW; 86DA1D67A37152A2 CRC64;
CC -----
CC Query Match 29.2%; Score 59.5; DB 1; Length 660;
CC Best Local Similarity 57.1%; Pred. No. 3 4;
CC Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
CC -----
CC QY 13 MGTCHSSDQPRGVPEPRCP 33
CC DB 1 MGTPCQARGPR-TTLPHPCP 20
CC -----
CC RESULT 4
CC NT02 HUMAN STANDARD; PRT; 2471 AA.
CC AC Q04721; Q99734; Q9H240;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (hM2).
CC GN NOTCH2.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Brain;
CC RA Blaumüller C.M., Mann R.S.;
CC RT "Complete human notch 2 (hM2) cDNA sequence.";
CC RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC [2]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Breast tumor;
CC RA Correa R.G., Canargo A.A., Moreira E.S., Simpson A.J.G.;
CC RT "Human Notch2, a novel member of cell-fate determining NOTCH
CC family.";
CC RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC [3]
CC RP SEQUENCE OF 967-1229 FROM N.A.
CC RC TISSUE=T-cell;
CC RA Lenasson I., Devaux C., Meenard J.M.;
CC RT "Partial sequence of EGF-like repeat domain of human Notch2 mRNA.";
CC RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC [4]
CC RP SEQUENCE OF 1810-2447 FROM N.A.
CC RC TISSUE=Brain;
CC RX MEDLINE=93265135; PubMed=1303260;
CC RA Stifani S., Blaumüller C.M., Redhead N.J., Hill R.E.,
CC RA Artavanis-Tsakonas S.;
CC RT "Human homologs of a Drosophila enhancer of split gene product define
CC a novel family of nuclear proteins.";
CC RL Nat. Genet. 2:119-127(1992).
CC [5]
CC RP POST-TRANSLATIONAL PROCESSING.
CC RX MEDLINE=97386453; PubMed=9244302;
CC RA Blaumüller C.M., Qi H., Zagouras P., Artavanis-Tsakonas S.;

```

```

RT "Intracellular cleavage of Notch leads to a heterodimeric receptor on
RT the plasma membrane.";
RL Cell 90:281-291(1997).
RN [6]
RX IDENTIFICATION OF LIGANDS.
RA MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
RA "Human ligands of the Notch receptor.";
RT
RL J. Pathol. 154:785-794(1999).
CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.
CC Upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity).
CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-
CC terminal fragment N(EC) which are probably linked by disulfide
CC bonds (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus.
CC -1- TISSUE SPECIFICITY: Expressed in the brain, heart, kidney, lung,
CC skeletal muscle and liver.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-Golgi network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TW) and a N-terminal fragment N(EC). Following
CC ligand binding, it is cleaved by TNF-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane (By similarity).
CC -1- PTM: Phosphorylated (By similarity).
CC -1- SIMILARITY: Belongs to the NOTCH family.
CC -1- SIMILARITY: Contains 35 EGF-like domains.
CC -1- SIMILARITY: Contains 2 lin/Notch repeats.
CC -1- SIMILARITY: Contains 6 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF308601; AAA36377.2; -
CC DR EMBL; AF315356; AAG37073.1; -
CC DR EMBL; U77493; AAB19224.1; -
CC DR HSSP; P00740; 1EDM
CC DR Genew; HGNC:7882; NOTCH2.
CC MIM; 600275; -
CC DR InterPro; IPR002110; ANK.
CC DR InterPro; IPR000152; Asx_hydroxyl_S.
CC DR InterPro; IPR000742; EGF_2.
CC DR InterPro; IPR001881; EGF_Ca.
CC DR InterPro; IPR001438; EGF-II.
CC DR InterPro; IPR006209; EGF_like.
CC DR InterPro; IPR002049; Laminin_EGF.
CC DR InterPro; IPR008297; Notch.
CC DR InterPro; IPR008000; Notch_dom.
CC DR Pfam; PF00023; ank; 6.
CC DR Pfam; PF00008; EGF; 35.
CC DR Pfam; PF00066; notch; 2.
CC DR PIRSF; PIRSF002279; Notch; 1.
CC DR PRINTS; PR00010; EGFLOOD.
CC DR PRINTS; PR00011; EGFAMININ.
CC DR PRINTS; PR01452; NOTCH.
CC DR SMART; SM00248; ANK; 6.
CC DR SMART; SM00179; EGF_CA; 23.

```

DR SMART; SMO0004; NL; 2.  
 DR PROSITE; PS00297; ANK REP REGION; 1.  
 DR PROSITE; PS00088; ANK REPEAT; 4.  
 DR PROSITE; PS00010; ASX HYDROXYL; 22.  
 DR PROSITE; PS00022; EGF 1; 34.  
 DR PROSITE; PS01186; EGF 2; 29.  
 DR PROSITE; PS00026; EGF 3; 35.  
 DR PROSITE; PS01187; EGF CA; 22.  
 KW Receptor; Transcription regulation; Activator; Differentiation;  
 KW Developmental protein; Receptor; ANK repeat; EGF-like domain;  
 KW Transmembrane; Glycoprotein; Signal; Phosphorylation.  
 FT SIGNAL 1 25  
 FT CHAIN 26 2471  
 FT CHAIN 1666 2471  
 FT CHAIN 1697 2471  
 FT DOMAIN 26 1677  
 FT TRANSMEM 1678 1698  
 FT DOMAIN 1699 2471  
 FT DOMAIN 26 63  
 FT DOMAIN 64 102  
 FT DOMAIN 105 143  
 FT DOMAIN 144 180  
 FT DOMAIN 182 219  
 FT DOMAIN 221 258  
 FT DOMAIN 260 296  
 FT DOMAIN 298 336  
 FT DOMAIN 338 374  
 FT DOMAIN 375 413  
 FT DOMAIN 415 454  
 FT DOMAIN 456 492  
 FT DOMAIN 494 530  
 FT DOMAIN 532 568  
 FT DOMAIN 570 605  
 FT DOMAIN 607 643  
 FT DOMAIN 645 680  
 FT DOMAIN 682 718  
 FT DOMAIN 720 755  
 FT DOMAIN 757 793  
 FT DOMAIN 795 831  
 FT DOMAIN 833 871  
 FT DOMAIN 873 909  
 FT DOMAIN 911 947  
 FT DOMAIN 949 985  
 FT DOMAIN 987 1023  
 FT DOMAIN 1025 1061  
 FT DOMAIN 1063 1099  
 FT DOMAIN 1101 1147  
 FT DOMAIN 1149 1185  
 FT DOMAIN 1187 1223  
 FT DOMAIN 1225 1262  
 FT DOMAIN 1264 1302  
 FT DOMAIN 1304 1343  
 FT DOMAIN 1374 1412  
 FT REPEAT 1420 1456  
 FT REPEAT 1503 1535  
 FT REPEAT 1827 1871  
 FT REPEAT 1875 1905  
 FT REPEAT 1909 1939  
 FT REPEAT 1943 1972  
 FT REPEAT 1976 2005  
 FT REPEAT 2009 2038  
 FT DOMAIN 1645 1648  
 FT DOMAIN 1994 1997  
 FT DOMAIN 2426 2429  
 FT DISULFID 28 41  
 FT DISULFID 35 51  
 FT DISULFID 53 62  
 FT DISULFID 68 79  
 FT DISULFID 73 90  
 FT DISULFID 92 101  
 FT DISULFID 109 121

FT DISULFID 115 131 BY SIMILARITY.  
 FT DISULFID 133 142 BY SIMILARITY.  
 FT DISULFID 148 159 BY SIMILARITY.  
 FT DISULFID 153 168 BY SIMILARITY.  
 FT DISULFID 170 179 BY SIMILARITY.  
 FT DISULFID 186 198 BY SIMILARITY.  
 FT DISULFID 192 207 BY SIMILARITY.  
 FT DISULFID 209 218 BY SIMILARITY.  
 FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 230 246 BY SIMILARITY.  
 FT DISULFID 248 257 BY SIMILARITY.  
 Query Match 28.7%; Score 58.5; DB 1; Length 2471;  
 Best Local Similarity 42.3%; Pred. No. 16;  
 Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;  
 Qy 7 GKCHCHMGTHC-HSSDGPGRGVPEPR 31  
 Db 1347 GQVCKRGQCVHTASGPRCFPSPR 1372  
 RESULT 5  
 SMRL MOUSE  
 ID SMRL\_MOUSE STANDARD; PRT; 147 AA.  
 AC Q61900;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Submaxillary gland androgen regulated protein 1 precursor (Salivary protein MSGI).  
 DE SMRL OR MSGI.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN-BALB/c; TISSUE-Submaxillary gland;  
 RX MEDLINE=94252564; PubMed=8194749;  
 RA Trautman R.D., Collins F.S., Senorale-Pose M., Rougeon F.;  
 RA "Three novel SMRL-related cDNAs characterized in the submaxillary gland of mice show extensive evolutionary divergence in the protein coding region."  
 RL Gene 142:175-182(1994).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE-Salivary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: May play a role in protection or detoxification.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Secreted into saliva by submaxillary gland.  
 CC -!- SIMILARITY: BELONGS TO THE PROL1/PROL3 FAMILY.  
 -----



Matches 10; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 4 YQKQCH--CHMGTHCSS 20  
 DB 253 YQKQCH--CHMGTHCSS 271

## RESULT 7

ID THYG HUMAN STANDARD; PRU; 2768 AA.  
 AC P0226; Q15274; Q43899; Q15593; Q9NVR2; Q9UMZ0; Q9UNY3;  
 DT 21-JUL-1986 (Rel. 04, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Thyroglobulin precursor.  
 GN TG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS ASP-604; ASP-653; GLN-985 DEL;  
 RP TYR-1043; THR-1059; ARG-1437; HIS-1463; THR-1936; GLU-2091; LEU-2149;  
 RP ARG-2170 AND HIS-2242.  
 EX MEDLINE=87246630; PubMed=3595599;  
 RA Malthiery Y., Lissitzky S.;  
 RT "Primary structure of human thyroglobulin deduced from the sequence  
 of its 8448-base complementary DNA.";  
 RL Eur. J. Biochem. 165:491-498(1987).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Thyroid;  
 RX MEDLINE=97329813; PubMed=9186272;  
 RA van de Graaf S.A.R., Pauw E., de Vijlder J.J.M., Ris-Stalpers C.;  
 RT "The revised 8307 base pair coding sequence of human thyroglobulin  
 transiently expressed in eukaryotic cells.";  
 RL Eur. J. Endocrinol. 136:508-515(1997).  
 [3]  
 RP SEQUENCE FROM N.A., AND VARIANT GLU-515.  
 RA Blechschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,  
 RA de Jonge R., Schilhabel M.B., Menzel U., Dette M.D., Baumgart C.,  
 RA Jahn N., Rosenthal A.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE OF 1-730 FROM N.A., AND VARIANTS ASP-604 AND ASP-653.  
 RX MEDLINE=85127024; PubMed=3971976;  
 RA Malthiery Y., Lissitzky S.;  
 RT "Sequence of the 5'-end quarter of the human-thyroglobulin messenger  
 ribonucleic acid and of its deduced amino-acid sequence.";  
 RL Eur. J. Biochem. 147:53-58(1985).  
 [5]  
 RP SEQUENCE OF 1-415; 640-737 AND 880-1000 FROM N.A., AND VARIANT  
 ALA-734.  
 RX MEDLINE=88062712; PubMed=3681978;  
 RA Parma J., Christophe D., Pohl V., Vassart G.;  
 RT "Structural organization of the 5' region of the thyroglobulin gene.  
 Evidence for intron loss and 'exonization' during evolution.";  
 RL J. Mol. Biol. 196:769-779(1987).  
 [6]  
 RP SEQUENCE OF 1-25 FROM N.A.  
 RX MEDLINE=85269632; PubMed=2991855;  
 RA Christophe D., Cabrer B., Bacolla A., Targovnik H.M., Pohl V.,  
 RA Vassart G.;  
 RT "An unusually long poly(purine)-poly(pyrimidine) sequence is located  
 upstream from the human thyroglobulin gene.";  
 RL Nucleic Acids Res. 13:5127-5144(1985).  
 [7]  
 RP SEQUENCE OF 1002-1566 FROM N.A.  
 RX PubMed=11124863;  
 RA Moya C.M., Mendive F.M., Rivolta C.M., Vassart G., Targovnik H.M.;  
 RT "Genomic organization of the 5' region of the human thyroglobulin  
 gene.";  
 RL Eur. J. Endocrinol. 143:789-798(2000).

RN [8]  
 RP SEQUENCE OF 1645-2768 FROM N.A., AND VARIANTS LEU-2149 AND ARG-2170.  
 RX MEDLINE=99452200; PubMed=10524569;  
 RA Mendive F.M., Rivolta C.M., Vassart G., Targovnik H.M.;  
 RT "Genomic organization of the 3' region of the human thyroglobulin  
 gene.";  
 RL Thyroid 9:903-912(1999).  
 [9]  
 RN PRESENCE OF A 11TH TYROGLOBULIN TYPE-I REPEAT.  
 RP MEDLINE=96390872; PubMed=8797845;  
 RA Molina F., Bouanani M., Pau B., Granier C.;  
 RT "Characterization of the type-1 repeat from thyroglobulin, a  
 cysteine-rich module found in proteins from different families.";  
 RL Eur. J. Biochem. 240:125-133(1996).  
 [10]  
 RN IODINATION SITES.  
 RP MEDLINE=89340430; PubMed=2760035;  
 RA Lamas L., Anderson P.C., Fox J.W., Dunn J.T.;  
 RT "Consensus sequences for early iodination and hormonogenesis in human  
 thyroglobulin.";  
 RL J. Biol. Chem. 264:13541-13545(1989).  
 [11]  
 RN SULFATION.  
 RP MEDLINE=99382264; PubMed=10448091;  
 RA Nlend M.-C., Cauvi D., Venot N., Chabaud O.;  
 RT "Sulfated tyrosines of thyroglobulin are involved in thyroid hormone  
 synthesis.";  
 RL Biochem. Biophys. Res. Commun. 262:193-197(1999).  
 [12]  
 RN VARIANT GOITER HIS-870.  
 RP MEDLINE=93164775; PubMed=8094490;  
 RA Corral J., Martin C., Perez R., Sanchez I., Mories M.T.,  
 RA San Millan J.L., Miralles J.M., Gonzalez-Sarmiento R.;  
 RT "Thyroglobulin gene point mutation associated with non-endemic simple  
 goitre.";  
 RL Lancet 341:462-464(1993).  
 [13]  
 RN VARIANTS GOITER ARG-1264 AND SER-1996, AND VARIANTS HIS-135; ASP-604;  
 ASP-653; ALA-734; GLU-830; GLN-985 DEL; VAL-1028; TYR-1043; THR-1059;  
 ASP-1312; ARG-1437; HIS-1463; ASN-1838; THR-1936; TRP-1999; GLU-2091;  
 LEU-2149; ARG-2170; HIS-2242; ARG-2501 AND GLN-2530.  
 RA Hishinuma A., Takamatsu J., Ohyama Y., Yokozawa T., Kanno Y., Kuma K.,  
 RA Yoshida S., Matsuura N., Ieiri T.;  
 RT "Two novel cysteine substitutions (C1263R and C1995S) of thyroglobulin  
 cause a defect in intracellular transport of thyroglobulin in  
 patients with congenital goiter and the variant type of adenomatous  
 goiter.";  
 RL J. Clin. Endocrinol. Metab. 84:1438-1444(1999).  
 CC -1- FUNCTION: Precursor of the iodinated thyroid hormones thyroxine  
 (T4) and triiodothyronine (T3).  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Thyroid gland specific.  
 CC -1- PM: Sulfated.  
 CC -1- DISEASE: Defects in TG are a cause of some forms of goiter  
 [MIM:188450, 274900]. Goiter is an enlargement of the thyroid  
 gland. This is sometimes linked to hypothyroidism.  
 CC -1- SIMILARITY: Belongs to the type-B carboxylesterase/lipase family.  
 CC -1- SIMILARITY: Contains 11 thyroglobulin type-I domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X05615; AAC29104.1; -;  
 CC EMBL; U93033; AAC51924.1; -;  
 CC EMBL; AF230667; -; NOT\_ANNOTATED\_CDS.  
 CC EMBL; AF235100; -; NOT\_ANNOTATED\_CDS.

DR	EMBL; AF230656; -; NOT ANNOTATED CDS.	
DR	EMBL; AF305872; -; NOT ANNOTATED CDS.	
DR	EMBL; X02154; CAA26089.1; -	
DR	EMBL; X06059; CAA29454.1; -	
DR	EMBL; X06061; CAA29454.1; JOINED.	
DR	EMBL; X06060; CAA29454.1; JOINED.	
DR	EMBL; X06062; CAA29454.1; JOINED.	
DR	EMBL; X06063; CAA29454.1; JOINED.	
DR	EMBL; X06064; CAA29454.1; JOINED.	
DR	EMBL; X06065; CAA29454.1; JOINED.	
DR	EMBL; X06066; CAA29454.1; JOINED.	
DR	EMBL; X06067; CAA29455.1; -	
DR	EMBL; X06068; CAA29455.1; JOINED.	
DR	EMBL; X06069; CAA29456.1; -	
DR	EMBL; X06070; CAA29456.1; JOINED.	
DR	EMBL; X02749; CAA26527.1; -	
DR	EMBL; AF170489; AAD51647.1; -	
DR	EMBL; AF170486; AAD51647.1; JOINED.	
DR	EMBL; AF170487; AAD51647.1; JOINED.	
DR	EMBL; AF170488; AAD51647.1; JOINED.	
DR	EMBL; AF105687; AAC95473.1; -	
DR	EMBL; AF105681; AAC95473.1; JOINED.	
DR	EMBL; AF105682; AAC95473.1; JOINED.	
DR	EMBL; AF105683; AAC95473.1; JOINED.	
DR	EMBL; AF105684; AAC95473.1; JOINED.	
DR	EMBL; AF105685; AAC95473.1; JOINED.	
DR	EMBL; AF105686; AAC95473.1; JOINED.	
DR	EMBL; AF080484; AAD50912.2; -	
DR	EMBL; AF169654; AAD50912.2; JOINED.	
DR	EMBL; AF169655; AAD50912.2; JOINED.	
DR	EMBL; AF169656; AAD50912.2; JOINED.	
DR	EMBL; AF169657; AAD50912.2; JOINED.	
DR	EMBL; AF169658; AAD50912.2; JOINED.	
DR	EMBL; AF169659; AAD50912.2; JOINED.	
DR	EMBL; AF169661; AAD50912.2; JOINED.	
DR	EMBL; AF169662; AAD50912.2; JOINED.	
DR	EMBL; AF169663; AAD50912.2; JOINED.	
DR	EMBL; AF169664; AAD50912.2; JOINED.	
DR	EMBL; AF080472; AAD50912.2; JOINED.	
DR	EMBL; AF080473; AAD50912.2; JOINED.	
DR	EMBL; AF080474; AAD50912.2; JOINED.	
DR	EMBL; AF080475; AAD50912.2; JOINED.	
DR	EMBL; AF080476; AAD50912.2; JOINED.	
DR	EMBL; AF080477; AAD50912.2; JOINED.	
DR	EMBL; AF080478; AAD50912.2; JOINED.	
DR	EMBL; AF080479; AAD50912.2; JOINED.	
DR	EMBL; AF080480; AAD50912.2; JOINED.	
DR	EMBL; AF080481; AAD50912.2; JOINED.	
DR	EMBL; AF080482; AAD50912.2; JOINED.	
DR	EMBL; AF080483; AAD50912.2; JOINED.	
DR	PIR; A59110; UIHU.	
DR	HSP; P21836; IMAA.	
DR	Genew; HGNC.11764; TG.	
DR	MIM; 188450; -	
DR	MIM; 274900; -	
DR	InterPro; IPR002018; CarboxypeptidaseB.	
DR	InterPro; IPR000716; Thyroglobulin_1.	
DR	Pfam; PF00135; Coesterase; 1.	
DR	Pfam; PF00086; thyroglobulin_1; 9.	
DR	SMART; SM00211; TY; 10.	
DR	PROSITE; PS00484; THYROGLOBULIN_1; 9.	
DR	PROSITE; PS00941; CARBOXYLPEPTIDASE_B_2; 1.	
KW	Glycoprotein; Repeat; Thyroid hormones; Thyroid hormones biosynthesis;	
Query Match 26.2%; Score 53.5; DB 1; Length 2768;		
Best Local Similarity 31.2%; Pred. No. 77;		
Matches 15; Conservative 2; Mismatches 12; Indels 19; Gaps 3;		
Qy	5 YQKFC-----HCHMGT-HCHSSDPRGVPE-----PRCP 33	
Db	1027 YMQCDAFGSEWPEVQCHAGTCHWCVDKGFIPGSLTARSLOIQCP 1074	

RESULT 8

ID	ITB6_CAVPO	STANDARD;	PRT;	577 AA.
AC	P18563;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Integrin beta-6 (fragment).			
GN	ITGB6.			
OS	Cavia porcellus (Guinea pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.			
OX	NCBI_TaxID=10141;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Hartley;			
RX	MEDLINE=90307659; PubMed=2365683;			
RA	Sheppard D., Rozzo C., Starr L., Quaranta V., Erle D.J., Pytela R.;			
RT	"Complete amino acid sequence of a novel integrin beta subunit (beta			
RT	6) identified in epithelial cells using the polymerase chain			
RT	reaction."			
RL	J. Biol. Chem. 265:11502-11507(1990).			
CC	-1- FUNCTION: INTEGRIN ALPHA-V/BETA-6 IS A RECEPTOR FOR FIBRONECTIN			
CC	AND CYTOTACTIN. IT RECOGNIZES THE SEQUENCE R-G-D IT ITS LIGANDS.			
CC	-1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Beta-6			
CC	associates with alpha-V. Interacts with FLNB (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- SIMILARITY: Belongs to the integrin beta chain family.			
CC	-1- SIMILARITY: Contains 2 WFPA-like domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
EMBL; M35197; AAA37043.1; -				
EMBL; A26611; CAA01833.1; -				
PIR; B37057; B37057.				
HSP; P05106; IJY2.				
InterPro; IPR006209; EGF_like.				
InterPro; IPR002369; Integrin_B.				
InterPro; IPR001169; Integrin_beta_C.				
Pfam; PF00362; integrin_B; 1.				
PRINTS; PR01186; INTEGRINB.				
ProDom; PD001811; Integrin_B; 1.				
SMART; SM00187; INB; 1.				
PROSITE; PS00243; INTEGRIN_BETA; 2.				
PROSITE; PS00022; EGF_1; UNKNOWN_1.				
PROSITE; PS01186; EGF_2; UNKNOWN_1.				
Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;				
Repeat.				
KW	NON TER	1		
FT	DOMAIN	<1	566	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	567	>577	POTENTIAL.
FT	DOMAIN	<1	230	WFPA-LIKE.
FT	DOMAIN	315	478	4 CYSTEINE-RICH TANDEM REPEATS.
FT	REPEAT	315	360	I.
FT	REPEAT	361	402	II.
FT	REPEAT	403	441	III.
FT	REPEAT	442	478	IV.
FT	DISULFID	56	63	BY SIMILARITY.
FT	DISULFID	111	152	BY SIMILARITY.
FT	DISULFID	253	265	BY SIMILARITY.
FT	DISULFID	285	529	BY SIMILARITY.
FT	DISULFID	311	315	BY SIMILARITY.
FT	DISULFID	326	338	BY SIMILARITY.
FT	DISULFID	335	370	BY SIMILARITY.
FT	DISULFID	340	349	BY SIMILARITY.
FT	DISULFID	351	361	BY SIMILARITY.
FT	DISULFID	376	381	BY SIMILARITY.

```

FT DISULFID 378 411 BY SIMILARITY.
FT DISULFID 383 396 BY SIMILARITY.
FT DISULFID 398 403 BY SIMILARITY.
FT DISULFID 417 422 BY SIMILARITY.
FT DISULFID 419 450 BY SIMILARITY.
FT DISULFID 424 433 BY SIMILARITY.
FT DISULFID 435 442 BY SIMILARITY.
FT DISULFID 456 461 BY SIMILARITY.
FT DISULFID 458 504 BY SIMILARITY.
FT DISULFID 463 473 BY SIMILARITY.
FT DISULFID 476 479 BY SIMILARITY.
FT DISULFID 483 492 BY SIMILARITY.
FT DISULFID 489 561 BY SIMILARITY.
FT DISULFID 508 537 BY SIMILARITY.
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 434 434 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 555 555 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 577 577
SQ SEQUENCE 577 AA; 62298 MW; B83B468C93EDCCF9 CRC64;

Query Match 25.7%; Score 52.5; DB 1; Length 577;
Best Local Similarity 34.2%; Pred. No. 23;
Matches 13; Conservative 2; Mismatches 14; Indels 9; Gaps 2;

QY 4 YQGKCHC---HMGTHCH-----SSDGRGVIPPRC 32
DB 333 YQGVACACNPGHMPHCEGCEGDTLSTDSKETPDHPC 370

RESULT 9
ID ZN44_HUMAN STANDARD; PRT; 589 AA.
AC P15621; P17018; Q9UL27;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 44 (Zinc finger protein K0X7) (Gonadotropin
DE inducible transcription repressor-2) (GIOT-2).
GN ZNF44 OR K0X7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RA Okada T.; Mizutani T.; Miyamoto K.;
RT "Identification and characterization of novel zinc finger proteins in
RT the human ovary."
RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 355-518 FROM N.A.
RA Ammendola S.; Ciliberto G.;
RN Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 365-420 FROM N.A.
RX MEDLINE=91145339; PubMed=2288909;
RA Thiesen H.-J.;
RT "Multiple genes encoding zinc finger domains are expressed in human T
RT cells."
RL New Biol. 2:363-374 (1990).
CC -1- FUNCTION: May be involved in transcriptional regulation.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -1- SIMILARITY: Contains 1 KRAB domain.

```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC EMBL; AB021642; BAA86988.1; -;  
CC EMBL; X16281; CAA34357.1; -;  
CC EMBL; X52338; CAA36564.1; -;  
CC PIR; D56409; S06779.  
CC HSP; P08048; 7ZNF.  
CC TRANSFAC; T04987; -;  
CC Genew; HGNC:13110; ZNF44.  
CC MIM; 194542; -;  
CC InterPro; IPR001909; KRAB.  
CC InterPro; IPR007087; Znf\_C2H2.  
CC InterPro; IPR007086; Znf\_C2H2\_sub.  
CC Pfam; PF01352; KRAB; 1.  
CC Pfam; PF00096; zf-C2H2; 15.  
CC PRINTS; PR00048; ZINC\_FINGER.  
CC ProDom; PD000003; Znf\_C2H2; 7.  
CC SMART; SM00349; KRAB; 1.  
CC SMART; SM00355; Znf\_C2H2; 15.  
CC PROSITE; PS00805; KRAB; 1.  
CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 13.  
CC PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 16.  
CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;  
CC Nuclear protein; Repeat.  
CC DOMAIN 4 90 KRAB.  
CC ZN\_FING 141 163 C2H2-TYPE (ATYPICAL).  
CC ZN\_FING 197 219 C2H2-TYPE.  
CC ZN\_FING 225 247 C2H2-TYPE.  
CC ZN\_FING 253 275 C2H2-TYPE.  
CC ZN\_FING 281 303 C2H2-TYPE.  
CC ZN\_FING 309 331 C2H2-TYPE.  
CC ZN\_FING 337 359 C2H2-TYPE.  
CC ZN\_FING 365 387 C2H2-TYPE.  
CC ZN\_FING 393 415 C2H2-TYPE.  
CC ZN\_FING 421 443 C2H2-TYPE.  
CC ZN\_FING 476 498 C2H2-TYPE.  
CC ZN\_FING 504 526 C2H2-TYPE.  
CC ZN\_FING 532 554 C2H2-TYPE.  
CC ZN\_FING 560 582 C2H2-TYPE.  
CC ZN\_FING 510 518 GKAFSRFSY -> AKPSVSVSVT (IN REF. 2).  
CC SEQUENCE 589 AA; 68102 MW; 24F3EA86617E7FDD CRC64;  
Query Match 25.7%; Score 52.5; DB 1; Length 589;  
Best Local Similarity 45.5%; Pred. No. 23;  
Matches 10; Conservative 2; Mismatches 5; Indels 5; Gaps 1;  
QY 7 GKCHCHMGTH-----CHSSDGP 23  
DB 315 GKAFCHLGSFQRHMHSGDGP 336  
RESULT 10  
ID ITB5\_HUMAN STANDARD; PRT; 799 AA.  
AC P18084;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Integrin beta-5 precursor.  
GN ITB5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A.  
RP SEQUENCE FROM N.A.

RC TISSUE=Thymic epithelium;  
RX MEDLINE=90282356; PubMed=2328726;  
RA Ramaswamy H., Hemler M.E.;  
RT "Cloning, primary structure and properties of a novel human integrin  
beta subunit";  
RL EMBO J. 9:1561-1568(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90319111; PubMed=2371275;  
RA Suzuki S., Huang Z.S., Tanihara H.;  
RT "Cloning of an integrin beta subunit exhibiting high homology with  
integrin beta 3 subunit";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:5354-5358(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91009141; PubMed=2211615;  
RA McLean J.W., Vestal D.J., Cheresch D.A., Bodary S.C.;  
RT "cDNA sequence of the human integrin beta 5 subunit";  
RL J. Biol. Chem. 265:17126-17131(1990).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1- FUNCTION: INTEGRIN ALPHA-V/BETA-5 IS A RECEPTOR FOR FIBRONECTIN.  
IT RECOGNIZES THE SEQUENCE R-G-D AT ITS LIGAND.  
CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-5  
ASSOCIATES WITH ALPHA-V.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: Belongs to the integrin beta chain family.  
CC -1- SIMILARITY: Contains 2 VNFA-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53002; CAA37188.1; --  
DR EMBL; M35011; AAA52707.1; --  
DR EMBL; J05633; AAA59183.1; --  
DR EMBL; BC006541; AAH06541.1; --  
DR F01; A38308; A38308.  
DR HSSP; P05106; 1JUV2.  
DR Genbank; HGNC:6160; ITGB5.  
DR MIM; 147561; --  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR002369; Integrin\_B.  
DR InterPro; IPR001169; Integrin\_beta\_C.  
DR InterPro; IPR003659; Flexin-like.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF00362; integrin\_B; 1.

DR PRINTS; PRO1186; INTEGRINB.  
DR ProDom; PD001811; Integrin\_B; 1.  
DR SMART; SM00187; INB; 1.  
DR SMART; SM00423; PSI; 1.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00243; INTEGRIN\_BETA; 2.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
DR PROSITE; PS01186; EGF\_2; UNKNOWN\_2.  
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycoprotein;  
KW Repeat; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 799 INTEGRIN BETA-5.  
FT DOMAIN 24 719 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 720 742 POTENTIAL.  
FT DOMAIN 743 799 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 136 378 VNFA-LIKE.  
FT DOMAIN 465 630  
FT REPEAT 465 512 I  
FT REPEAT 513 554 II.  
FT REPEAT 555 593 III.  
FT REPEAT 594 630 IV.  
FT DISULFID 28 463 BY SIMILARITY.  
FT DISULFID 36 46 BY SIMILARITY.  
FT DISULFID 39 75 BY SIMILARITY.  
FT DISULFID 49 64 BY SIMILARITY.  
FT DISULFID 202 211 BY SIMILARITY.  
FT DISULFID 259 300 BY SIMILARITY.  
FT DISULFID 401 413 BY SIMILARITY.  
FT DISULFID 433 682 BY SIMILARITY.  
FT DISULFID 461 465 BY SIMILARITY.  
FT DISULFID 476 487 BY SIMILARITY.  
FT DISULFID 484 522 BY SIMILARITY.  
FT DISULFID 489 498 BY SIMILARITY.  
FT DISULFID 500 513 BY SIMILARITY.  
FT DISULFID 528 533 BY SIMILARITY.  
FT DISULFID 530 563 BY SIMILARITY.  
FT DISULFID 535 548 BY SIMILARITY.  
FT DISULFID 550 555 BY SIMILARITY.  
FT DISULFID 569 574 BY SIMILARITY.  
FT DISULFID 571 602 BY SIMILARITY.  
FT DISULFID 576 585 BY SIMILARITY.  
FT DISULFID 587 594 BY SIMILARITY.  
FT DISULFID 608 613 BY SIMILARITY.  
FT DISULFID 610 625 BY SIMILARITY.  
FT DISULFID 628 631 BY SIMILARITY.  
FT DISULFID 635 644 BY SIMILARITY.  
FT DISULFID 641 714 BY SIMILARITY.  
FT DISULFID 661 690 BY SIMILARITY.  
FT CARBOHYD 347 347 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 460 460 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 477 477 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 505 505 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 552 552 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 654 654 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 705 705 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 193 193 T -> A (IN REF. 2).  
FT CONFLICT 645 645 L -> P (IN REF. 3).  
FT CONFLICT 790 792 MISSING (IN REF. 2).  
SQ SEQUENCE 799 AA; 88053 MW; D7E4727CA310512B CRC64;  
Query Match 25.7%; Score 52.5; DB 1; Length 799;  
Best Local Similarity 31.4%; Pred. No. 31;  
Matches 11; Conservative 4; Mismatches 17; Indels 3; Gaps 1;  
QY 1 STVYQGRKCHC---HMGTGCHSSDGRGVIPERPC 32  
DB 479 SGTVCGLCGSPCYLGLTRCECDGENQSVYQNL 513  
RESULT 11  
DMN\_HUMAN











```
FT DISULFID 956 965 BY SIMILARITY.
FT DISULFID 968 983 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 988 1007 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1022 1035 BY SIMILARITY.
FT DISULFID 1095 1107 BY SIMILARITY.
FT DISULFID 1097 1114 BY SIMILARITY.
FT DISULFID 1116 1125 BY SIMILARITY.
FT DISULFID 1128 1140 BY SIMILARITY.
FT DISULFID 1143 1155 BY SIMILARITY.
FT DISULFID 1145 1162 BY SIMILARITY.
FT DISULFID 1164 1173 BY SIMILARITY.
FT DISULFID 1176 1187 BY SIMILARITY.
FT DISULFID 1190 1190 INTERCHAIN (PROBABLE).
FT DISULFID 1193 1193 INTERCHAIN (PROBABLE).
FT DISULFID 1797 1797 INTERCHAIN (PROBABLE).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1249 1249 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1308 1308 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 914 914 R -> G (IN REF. 2).
FT CONFLICT 1179 1179 G -> A (IN REF. 2).
SQ SEQUENCE 1798 AA; 196079 MW; 9555CF5B24850CB7 CRC64;

Query Match 25.5%; Score 52; DB 1; Length 1798;
Best Local Similarity 23.9%; Pred. No. 78;
Matches 11; Conservative 4; Mismatches 13; Indels 19; Gaps 1;

QY 5 YQKCHCHM-----GTHCHSSDPRGVPEPC 32
DB 1110 FTQCHCRAGFGGRTCEQELHWGDPGLQCHACDCDSRGIDTPQC 1155

RESULT 14
HMRI RHIME
ID HMRI RHIME STANDARD; PRT; 147 AA.
AC P58378;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Heavy metal dependent transcription regulator 1.
GN HMRI1 OR HMR1 OR R20549 OR SNA1014
OS Rhizobium melliloti (Sinorhizobium melliloti).
OG Plasmid pSyaA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021.
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Guzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federpiei N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium melliloti pSyaA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
CC -!- FUNCTION: TRANSCRIPTIONAL REGULATOR INVOLVED IN ACID TOLERANCE.
CC -!- BINDS COPPER (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- DOMAIN: It contains a N-terminal DNA binding region and a C-
CC terminal metal binding region (By similarity).
CC -!- SIMILARITY: Contains 1 HTH merR-type DNA-binding domain.
CC -!- CAUTION: We have changed a potential AGG for Arg-1 as ATG for Met-
CC 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; AE007245; AAK65207.1; ALT_SEQ.
DR InterPro; IPR000551; HTH_MerR.
DR Pfam; PF00376; merR; 1.
DR PRINTS; PR00040; HTHMERR.
DR SMART; SM00422; HTH_MERR; 1.
DR PROSITE; PS00552; HTH_MERR_1; FALSE NEG.
DR PROSITE; PS00937; HTH_MERR_2; 1.
KW Transcription regulation; DNA-binding; Copper; Plasmid;
KW Complete proteome.
FT DOMAIN 1 70 HTH_MERR-TYPE.
FT DNA_BIND 3 22 H-T-H MOTIF (POTENTIAL).
SQ SEQUENCE 147 AA; 16415 MW; 3A447536A9C07A86 CRC64;

Query Match 25.2%; Score 51.5; DB 1; Length 147;
Best Local Similarity 36.4%; Pred. No. 8.1;
Matches 8; Conservative 2; Mismatches 5; Indels 7; Gaps 1;

QY 12 HMGTHCHSSDPRGVPEPC 33
DB 108 HLASHCHGDD-----RPECP 122

RESULT 15
WIFI BRARE
ID WIFI BRARE STANDARD; PRT; 378 AA.
AC Q9W6F9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE WNT inhibitory factor 1 precursor (WIF-1).
GN WIFI.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9215557; PubMed=10201374;
RA Hsieh J.-C., Kodjabachian L., Rebbert M.L., Rattner A.,
RA Smallwood P.M., Samos C.H., Nusse R., Dawid I.B., Nathans J.;
RT "A new secreted protein that binds to Wnt proteins and inhibits their
RT activities.";
RL Nature 398:431-436 (1999).
CC -!- FUNCTION: Binds to WNT proteins and inhibits their activities. May
CC be involved in mesoderm segmentation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Highly expressed in unsegmented paraxial
CC mesoderm.
CC -!- SIMILARITY: Contains 5 EGF-like domains.
CC -!- SIMILARITY: Contains 1 WIF domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; AF122925; AAD25405.1; -.
DR PIR; B59180; B59180.
DR HSPF; P00740; 1EDM.
DR ZFIN; ZDB-GENE-990712-17; wif1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; EGF.
```

DR InterPro; IPR002049; Laminin\_EGF.  
DR InterPro; IPR003306; WIF.  
DR Pfam; PF00008; EGF; 3.  
DR Pfam; PF02019; WIF; 1.  
DR PRINTS; PR00011; EGFLAMININ.  
DR SMART; SM00181; EGF; 5.  
DR SMART; SM00469; WIF; 1.  
DR PROSITE; PS00022; EGF\_1; 5.  
DR PROSITE; PS01186; EGF\_2; 4.  
DR PROSITE; PS50026; EGF\_3; 4.  
DR PROSITE; PS50814; WIF; 1.  
KW Repeat; EGF-like domain; Signal; Developmental protein;  
KW Wnt signaling pathway.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 378 WNT INHIBITORY FACTOR 1.  
FT DOMAIN 36 174 WIF.  
FT DOMAIN 175 206 EGF-LIKE 1.  
FT DOMAIN 207 238 EGF-LIKE 2.  
FT DOMAIN 239 270 EGF-LIKE 3.  
FT DOMAIN 271 302 EGF-LIKE 4.  
FT DOMAIN 303 334 EGF-LIKE 5.  
FT DISULFID 175 184 POTENTIAL.  
FT DISULFID 180 190 POTENTIAL.  
FT DISULFID 196 198 POTENTIAL.  
FT DISULFID 207 216 POTENTIAL.  
FT DISULFID 212 222 POTENTIAL.  
FT DISULFID 228 230 POTENTIAL.  
FT DISULFID 239 248 POTENTIAL.  
FT DISULFID 244 254 POTENTIAL.  
FT DISULFID 260 262 POTENTIAL.  
FT DISULFID 271 280 POTENTIAL.  
FT DISULFID 276 286 POTENTIAL.  
FT DISULFID 292 294 POTENTIAL.  
FT DISULFID 303 312 POTENTIAL.  
FT DISULFID 308 318 POTENTIAL.  
FT DISULFID 324 326 POTENTIAL.  
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 378 AA; 41312 MW; 42FE9F70D948D1D8 CRC64;

Query Match 25.2%; Score 51.5; DB 1; Length 378;  
Best Local Similarity 36.1%; Pred. No. 20;  
Matches 13; Conservative 1; Mismatches 13; Indels 9; Gaps 2;  
QY 6 QGKCHCHMGTH-----CHSSDGRGVIPER-RC 32  
Db 289 RNKCKCKGYHGLCSKAVCEPSCGAHGTCTVEPNC 324

Search completed: February 26, 2004, 12:10:52  
Job time : 5.80106 secs



RX MEDLINE=93208881; PubMed=7916655;  
RA Browning J.L., Ngan-ek A., Lawton P., DeMarinis J., Tizard R.,  
RA Chou E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.,  
RT "lymphotoxin-beta: A new member of the TNF family that forms a  
RT heteromeric complex with lymphotoxin on the cell surface.";  
RL Cell 72:847-856(1993).  
RN [5]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=86016093; PubMed=2995927;  
RX Nedwin G.E., Navlor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
RA Pennica D., Goeddel D.V., Gray P.W.,  
RA "Human lymphotoxin and tumor necrosis factor genes: structure,  
RT homology and chromosomal localization.";  
RL Nucleic Acids Res. 13:6361-6373(1985).  
RN [6]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=91086846; PubMed=1670638;  
RX Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
RA Riethmuller G., Weiss E.H.,  
RA "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
RT NcoI Polymorphism in the First Intron of the Human TNF-beta Gene  
RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
RT Level of TNF-beta Production.";  
RL J. Exp. Med. 173:209-219(1991).  
RN [7]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=91139175; PubMed=1671667;  
RX Abraham J.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.,  
RA "Haplotypic polymorphisms of the TNF gene.";  
RL Immunogenetics 33:50-53(1991).  
RN [8]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=94362679; PubMed=8081366;  
RX Albertella M.R., Campbell D.R.,  
RA "Characterization of a novel gene in the human major  
RT histocompatibility complex that encodes a potential new member of the  
RT I kappa B family of proteins.";  
RL Hum. Mol. Genet. 3:793-799(1994).  
RN [9]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=95324911; PubMed=7601445;  
RX Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
RA Van Zeveren A., Coppeters W., Van de Weghe A., Bouquet Y., Choy W.,  
RA Strominger J., Spies T.,  
RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
RT Nuclear RNA Helicase of the D-E-A-D Family.";  
RL Genomics 26:210-218(1995).  
RN [10]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=20132445; PubMed=10668961;  
RX Neville M.J., Campbell R.D.,  
RA "Alternative splicing of the LST-1 gene located in the major  
RT histocompatibility complex on human chromosome 6.";  
RL DNA Seq. 8:155-160(1997).  
RN [11]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=98035989; PubMed=9367684;  
RX de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
RA Weiss E.H.,  
RT "Complex expression pattern of the TNF region gene LST1 through  
RT differential regulation, initiation, and alternative splicing.";  
RL Genomics 45:591-600(1997).  
RN [12]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=98149985; PubMed=9480751;  
RX Shilina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
RA Kimura M., Inoko H.,  
RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
RT IRL and MICA genes at the centromeric end of the HLA class I  
RL Genomics 47:372-382(1998).

DR EMBL; Y14768; CAA75068.1; --  
DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR GO; GO:0006954; P:inflammatory response; NAS.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;  
  
Query Match 100.0%; Score 204; DB 4; Length 165;  
Best Local Similarity 100.0%; Pred. NO. 6.1e-21;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 STVYQKCHCHGTHCHSSDGRGVIPERCP 33  
|||||  
DB 133 STVYQKCHCHGTHCHSSDGRGVIPERCP 165  
  
RESULT 2  
O14932  
ID O14932 PRELIMINARY; PRT; 190 AA.  
AC O14932;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE IC7 precursor (Natural killer cell receptor).  
GN IC7 OR NKP30.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=96422187; PubMed=8824804;  
RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.,  
RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";  
RL Genomics 31:215-222(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RA Nalabolu S.R., Raghunathan A., Weissman S.M.,  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93272029; PubMed=8499947;  
RA Iris F., Bougueret L., Prieur S., Caterina D., Primas G., Perrot V.,  
RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.,  
RT "Dense Alu clustering and a potential new member of the NF kappa B  
RT family within a 90 kilobase HLA class III segment.";  
RL Nat. Genet. 3:137-145(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96215741; PubMed=8629302;  
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,  
RA Wallace A.F., Russell M.E.,  
RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
RT molecule expressed in transplanted human hearts.";  
RL Transplantation 61:1387-1392(1996).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96006565; PubMed=7590964;  
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierina H.,  
RA Weiss E.H.,  
RT "Cloning and genomic characterization of LST1: a new gene in the human  
RT TNF region.";  
RL Immunogenetics 42:315-322(1995).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93208881; PubMed=7916655;  
RA Browning J.L., Ngan-ek A., Lawton P., DeMarinis J., Tizard R.,

RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface."  
 RL Cell 72:847-856(1993).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization."  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production."  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene."  
 RL Immunogenetics 33:50-53(1991).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins."  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vaiman M.,  
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Stronger J., Spies T.;  
 RT "The BARI gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family."  
 RL Genomics 26:210-218(1995).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6."  
 RL DNA Seq. 8:155-160(1997).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinorzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing."  
 RL Genomics 45:591-600(1997).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shihina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Tanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IKB and MICA genes at the centromeric end of the HLA Class I  
 RT region."  
 RL Genomics 47:372-382(1998).  
 RN [15]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Lymphoid;  
 RA Blassoni R., Pessino A., Malaespina A.;  
 RT "NK-A1 activating NK receptor."  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Peripheral blood;  
 RA Sato M., Yabe T., Ohashi J., Tsuchiya N., Hanaoka K., Tokunaga K.,  
 RA Fuji T.;  
 RT "Identification of two novel single nucleotide polymorphisms in the  
 RT NKp30 gene in human natural killer cells."  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF031138; AAS86590.1; -  
 DR EMBL; Y14768; CAA75065.1; -  
 DR EMBL; AJ223153; CAB54004.1; -  
 DR EMBL; AB055881; BAB78472.1; -  
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW SIGNAL; Receptor.  
 FT SIGNAL 1 61 POTENTIAL.  
 FT CHAIN 62 190 1C7  
 SQ SEQUENCE 190 AA; 20640 MW; 1FAC919E20A6B18A CRC64;  
 Query Match 100.0%; Score 204; DB 4; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 7e-21;  
 Matches 33; Conservative 0; Mismatches 0; Gaps 0;  
 Qy 1 STVYQKCHCHMGTHCHSSDGRGVIPRCP 33  
 Db 158 STVYQKCHCHMGTHCHSSDGRGVIPRCP 190  
 RESULT 3  
 Q8MJ02 PRELIMINARY; PRT; 180 AA.  
 AC Q8MJ02;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NKp30.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LaBonte M.L., Miller J., Letvin N.L.;  
 RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification  
 RT of Nkp46SD and Nkp30S."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY035214; AAK63116.1; -  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00047; Ig; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;  
 Query Match 69.6%; Score 142; DB 6; Length 180;  
 Best Local Similarity 95.7%; Pred. No. 3.6e-12;  
 Matches 22; Conservative 1; Mismatches 0; Gaps 0;  
 Qy 1 STVYQKCHCHMGTHCHSSDGP 23  
 Db 158 STVYQKCHCHMGTHCHSSDGP 180

```
RESULT 4
Q8MJ01 ID Q8MJ01 PRELIMINARY; PRT; 151 AA.
AC Q8MJ01
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NKP30v1ED.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA LaBonte M.L., Miller J., Letvin N.L.;
RT "Molecular cloning of rhesus monkey NKP46 and identification
of NKP46SD and NKP30S."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RE EMBL: AY035216; AAK63118.1; -
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;

Query Match 58.3%; Score 119; DB 6; Length 151;
Best Local Similarity 94.7%; Pred. No. 5.3e-09;
Matches 18; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHS 19
Db 133 STLYYQKCHCHMGTHCHS 151

RESULT 5
Q95JB8 ID Q95JB8 PRELIMINARY; PRT; 176 AA.
AC Q95JB8
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NKP30 (NKP30v1).
GN NCR.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541, 9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Rizzi M., Biassoni R.;
RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in
Macaca fascicularis lymphoid cells."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RE EMBL: AJ278389; CA41081.1; -
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;

Query Match 32.8%; Score 67; DB 11; Length 192;
Best Local Similarity 71.4%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCHCHMG 179

RESULT 6
Q8CG11 ID Q8CG11 PRELIMINARY; PRT; 192 AA.
AC Q8CG11
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NK receptor 1c7 precursor.
GN 1c7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor 1c7."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RE EMBL: AJ430419; CAD23067.2; -
DR EMBL; AJ430419; CAD23067.2; JOINED.
GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SMO0409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Signal; Receptor.
FT SIGNAL 1 19 POTENTIAL.
SQ SEQUENCE 192 AA; 20470 MW; 439AD7A3AFBE6DC0 CRC64;

Query Match 32.8%; Score 67; DB 11; Length 192;
Best Local Similarity 71.4%; Pred. No. 0.14;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCHCHMG 179

RESULT 7
Q8OWM8 ID Q8OWM8 PRELIMINARY; PRT; 192 AA.
AC Q8OWM8
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NKP30.
GN NKP30.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22168131; PubMed=12180816;
RX Hsieh C.L., Obata H., Ogura Y., Martinez O.M., Krams S.M.;
RT "NK cells and transplantation."
RL Transpl. Immunol. 9:111-114(2002).
RN [2]
RP SEQUENCE FROM N.A.
SQ SEQUENCE 192 AA; 20470 MW; 439AD7A3AFBE6DC0 CRC64;
```



```

RA Hsieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RA Martinez O.M., Krams S.M.;
RT "Identification, Cloning, and Characterization of a Novel Rat NK
RT Receptor, rNKP30: a Molecule Expressed in Liver Allografts.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL: AY273824; AAP13457.1; -.
DR InterPro: IPR003599; I9.
DR InterPro: IPR007110; I9-like.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
SQ SEQUENCE 192 AA; 20498 MW; 7FDF58B245C52377 CRC64;

Query Match
Best Local Similarity 32.8%; Score 67; DB 11; Length 192;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCLCHVG 179

RESULT 8
Q8CFD9 PRELIMINARY; PRT; 192 AA.
AC Q8CFD9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE I27 protein precursor.
GN I27.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=NK cell;
RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor I27.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL: A430418; CAD23066.1; -.
DR InterPro: IPR003599; I9.
DR InterPro: IPR007110; I9-like.
DR Pfam: PF00047; IG; 1.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG_LIKE; 1.
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 192 I27 PROTEIN.
SQ SEQUENCE 192 AA; 20500 MW; 7FDD5AB252D239C7 CRC64;

Query Match
Best Local Similarity 32.8%; Score 67; DB 11; Length 192;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMG 14
Db 166 SVIYQKCLCHVG 179

RESULT 9
Q7WQL1 PRELIMINARY; PRT; 665 AA.
AC Q7WQL1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hydatoin utilization protein B.
GN HYB OR BB0319.
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.

```

```

OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RS50 / ATCC BAA-588;
MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Mungall K.L.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jørgensen K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR ENBL: BX640437; CAE30817.1; -.
KW Complete proteome.
SQ SEQUENCE 665 AA; 72398 MW; 5B244ED30DDA2762 CRC64;

```

```

Query Match
Best Local Similarity 31.1%; Score 63.5; DB 16; Length 665;
Matches 12; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 VYQKCHCHMGTHCHSSDGRGVIP 28
Db 108 VYQGMVCFVGGHINTD-MGGAVP 132

```

```

RESULT 10
Q7WING PRELIMINARY; PRT; 665 AA.
AC Q7WING;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hydatoin utilization protein B.
GN HYB OR BFP0316.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Mungall K.L.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Quail M.A.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jørgensen K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR ENBL: BX640423; CAE40057.1; -.
KW Complete proteome.
SQ SEQUENCE 665 AA; 72497 MW; 5B2BEFC60DDE39D2 CRC64;

```

```

Query Match
Best Local Similarity 31.1%; Score 63.5; DB 16; Length 665;
Matches 12; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 VYQKCHCHMGTHCHSSDGRGVIP 28
Db 108 VYQGMVCFVGGHINTD-MGGAVP 132

```

**RESULT 11**

ID QVZW9 PRELIMINARY; PRT; 665 AA.

AC QVZW9;

DT 01-OCT-2003 (TEMBLrel. 25, Created)

DR 01-OCT-2003 (TEMBLrel. 25, Last sequence update)

DY 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE Hydatonin utilization protein B.

GN HYD OR BP0754.

OS Bordetella pertussis.

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

CC Alcaligenaceae; Bordetella.

OX NCBI\_TaxID=520;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Tchana I / ATCC BAA-589 / NCTC 13251;

RX MEDLINE=22827934; PubMed=12910271;

RA Parkhill J., Sebathia M.T.G., Churcher C.M., Bentley S.D., Mungall K.L., Harris D.E., Holden M.T.G., Churchev C.M., Quail M.A., Cardeno-Tarraga A.M., Temple L., James K., Basham D., Ashton R., Chervach I., Achtman M., Atkin R., Baker S., Basham D., Basson N., Chervach I., Chillingworth T., Collins M., Cronin A., Davis P., Doggett J., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Hornby T., Leather S., Moule S., Norbertczak H., O'Neill S., Ormond D., Price C., Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Stevens K., Unwin L., Whitehead S., Barrrell B.G., Maskell D.J.; "Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica"; Nat. Genet. 35:32-40(2003);

DB ENBL: BX640413; CAB41060.1; .

KW Complete Proteome.

RZ SEQUENCE 665 AA; 72395 MW; F27204349A190508 CRC64;

SQ

Query Match 31.1%; Score 63.5; DB 16; Length 665;  
Best Local Similarity 46.2%; Pred.No.1.5;  
Matches 12; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 VYYGKCHCHNGTHCSSDGPRGVIP 20  
||||| : : : : :  
DB 108 VYVGGEWCVFVGHHTND-MGGAVP 132

**RESULT 12**

ID Q97271 PRELIMINARY; PRT; 485 AA.

AC Q97271

DT 01-MAY-1999 (TEMBLrel. 10, Created)

DR 01-MAR-2003 (TEMBLrel. 23, Last sequence update)

DY 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

GN MAL35.20.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI\_TaxID=36329;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=9376085; PubMed=10448855;

RX Bowman S., Lawson D., Basham D., Brown D., Chillingworth T., Churcher C.M., Craig A., Davies R.M., Devlin K., Feltwell T., Gentles S., Williams R., Hamlin N., Harris D., Holroyd S., Hornby T., Horrocks P., Jagels K., Jassal B., Kyes S., McLean J., Moule S., Mungall K., Murphy L., Oliver K., Quail M.A., Rajandram M.-A., Rutter S., Skelton J., Squares R., Stevens K., Sulston J.B., Whitehead S., Woodward J.R., Newbold C., Barrrell B.G.; "The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum"; Nature 400:532-538 (1999).

RZ SEQUENCE FROM N.A.

RX MEDLINE=932255708; PubMed=12368867;

RA Hall N., Pain A., Berrihan M., Churcher C., Harris B., Harris D.,

Mungall K., Bowman S., Atkin R., Baker S., Basham D., Brooks K., Buckee C.O., Burrows C., Chervach I., Chillingworth C., Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C., Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J., Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z., Harper D., Hauser H., Hornby T., Holroyd S., Horrocks P., Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A., Knights A., Konfortov B., Kyes S., Latke N., Lawson D., Lemard N., Line A., Maddison M., Mclean J., Mooney P., Moule S., Murphy L., Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E., Rajandram M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M., Seeger K., Sharp S., Smith R., Squares R., Squares R., Stevens K., Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J., Sulston J.B., Craig A., Newbold C., Barrrell B.G.; "Sequence of plasmodium falciparum chromosomes 1, 3-9 and 13."; Nature 419:527-531(2002).

RL Nature 419:527-531(2002).

DR EMBL: AL034556; CAB38982.3; .

GO GO:0005489; F:electron transporter activity; IEA.

GO GO:0006118; P:electron transport; IEA.

DR InterPro; IPRO00345; Cytochrome\_C\_1.

PROSITE; PS00190; CYTOCHROME\_C\_1.

KW Hypothetical protein.

SQ SEQUENCE 485 AA; 57072 MW; 59C4467FC026CA72 CRC64;

Query Match 29.9%; Score 61; DB 5; Length 485;  
Best Local Similarity 25.0%; Pred.No.2.5;  
Matches 11; Conservative 10; Mismatches 7; Indels 16; Gaps

QY 4 YVYGKCH-----CHMGTHC-----HSSDGRGVPEPR 31  
:::| | | | |  
DB 156 HHKRGCSCKVEWTGCHLSRFRCDHQSHVPETQRVVFPNQ 199

**RESULT 13**

ID Q95667 PRELIMINARY; PRT; 176 AA.

AC Q95667;

DT 01-MAY-1999 (TEMBLrel. 10, Created)

DY 01-MAY-1999 (TEMBLrel. 10, Last sequence update)

DY 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE 1C7e.

GN 1C7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=93272029; PubMed=8499947;

RX Iris F., Bougueret L., Priour S., Caterina D., Primas G., Perrot V., Jurka J., Rodriguez-Tome P., Clavetier J., Cohen D., Daussset J.; "Dense Alle clustering and a potential new member of the NFkappaB family within a 90 kilobase HLA class III segment."; Nat. Genet. 3:137-145(1993).

RL [2]

RP SEQUENCE FROM N.A.

RC MEDLINE=96215741; PubMed=8629302;

RX Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J., Wallace A.F., Russell M.E.; "Allograft inflammatory factory-1. A cytokine-responsive macrophage molecule expressed in transplanted human hearts.";

RL Transplantation 61:1387-1392(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC MEDLINE=96006565; PubMed=7590964;

RX Holzinger I., de Baey G., Messer G., Kick G., Zwierzina H., Weiss E.H.; "Cloning and genomic characterization of LST1: a new gene in the human TNF region.";

RL Immunogenetics 42:315-322 (1995).

RZ SEQUENCE FROM N.A.

RX MEDLINE=93208881; PubMed=7916655;

```

DR GO: GO:0003793; F:defense/immunity protein activity; NAS.
DR GO: GO:0006955; P:immune response; NAS.
DR GO: GO:0006954; P:inflammatory response; NAS.
DR InterPro: IPR003599; IG.
DR InterPro: IPR007110; IG-like.
DR Pfam: PF00047; IG; 1.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS00835; IG-LIKE; 1.
SQ SEQUENCE 176 AA; 18749 MW; 162BB775DA2BCD35 CRC64;

Query Match          29.4%; Score 60; DB 4; Length 176;
Best Local Similarity 43.2%; Pred. No. 1.3;
Matches 16; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 1 STVYQKCHCHMGTHCHSSDGR-----GVIEPRCP 33
   |||||
DB 133 STVYQK-----CLTWKGRQLPAVVPAFLPP 161

RESULT 14
ID O14931 PRELIMINARY; PRT; 201 AA.
AC O14931
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 1C7 precursor (NCR3 protein).
GN DE GN
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
CB [1]
SEQUENCE FROM N.A.
TI TISSUE-Spleen;
MD MEDLINE=96422187; PubMed=8824804;
RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;
RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."
RL Genomics 31:215-222(1996).
[2]
SEQUENCE FROM N.A.
TI TISSUE-Spleen;
RA Nalabolu S.R., Raghunathan A., Weissman S.M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBSJ databases.
[3]
SEQUENCE FROM N.A.
RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
RA Laskey S., Hood L.;
RT "Sequence of the human major histocompatibility complex class III
RL region."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBSJ databases.
[4]
SEQUENCE FROM N.A.
RA MEDLINE=93272029; PubMed=8499947;
RA Iris F., Bouqueret L., Prieur S., Caterina D., Primas G., Perrot
RA Jurka J., Rodriguez-Tome P., Clavier J., Cohen D., Dausset J.;
RA "Dense Alu clustering and a potential new member of the NF kappa B
RT family within a 90 kilobase HLA class III segment."
RL Nat. Genet. 3:137-145(1993).
[5]
SEQUENCE FROM N.A.
RA MEDLINE=96215741; PubMed=8629302;
RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,
RA Wallace A.F., Russell M.R.;
RA "Allergic inflammatory factory-1. A cytokine-responsive macrophage
RT molecule expressed in transplanted human hearts."
RL Transplantation 61:1387-1392(1996).
[6]
SEQUENCE FROM N.A.
RA MEDLINE=96006565; PubMed=7590964;
RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,
RA Weiss E.H.;

```



```

RN SEQUENCE FROM N.A.
RP Strausberg R.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RL ENBL; BC026051; AAH26051.1; -.
DR MGD; MGI:99916; Lamb2.
DR GO; GO:0005578; C:extracellular matrix; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004623; F:phospholipase A2 activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0016042; P:lipid catabolism; IEA.
DR InterPro; IPR005797; Cytb_b6_N.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR02049; Laminin_EGF.
DR InterPro; IPR008211; LamNT.
DR InterPro; IPR01211; PhospholipaseA2.
DR Pfam; PF00053; laminin_EGF; 13.
DR Pfam; PF00055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00180; EGF_Lam; 13.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00152; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00022; EGF_1; 10.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 12.
DR PROSITE; PS00119; PA2_ASP; 1.
KW Laminin EGF-like domain.
SQ SEQUENCE 1799 AA; 196578 MW; 37CA24B9CDA0791F CRC64;

```

```

Query Match          29.4%; Score 60; DB 11; Length 1799;
Best Local Similarity 29.8%; Pred. No. 13;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

```

```

QY      5 YGKCHCHM-----GTHCHSSD-GPGVPEPRC 32
Db     1111 FTGQCHAGFGGRTSCQELYWGDPLQCRACDCDPRG-IDKPOC 1156

```

```

Search completed: February 26, 2004, 12:13:48
Job time : 17.0928 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 20.4828 Seconds  
(without alignments)  
455.215 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHGTHCHSSDPRGVPEPRCP 33

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_25Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003s.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	33	Aae02773	Human Nkp
2	204	100.0	190	Aay06401	Human B-c
3	204	100.0	190	Aae02769	Human Nkp
4	61.5	30.1	82	Aag35071	Arabidops
5	61.5	30.1	82	Aag11882	Arabidops
6	61.5	30.1	636	Aab66269	Rat TANGO
7	60	29.4	66	Aaol0665	Human pol
8	60	29.4	201	Aay06403	Human B-c
9	60	29.4	1799	Aam50359	Mouse lam
10	60	29.4	1801	Aaw50895	Rat lamin
11	60	29.4	1801	Ad60383	Rat Prote
12	59	28.9	57	Aau43863	Propionib
13	59	28.9	57	Abm40382	Propionib
14	59	28.9	2813	Aaw54347	Canine vo
15	59	28.9	2813	Aay70557	Canine vo
16	58.5	28.7	77	Ab597268	Novel hum
17	58.5	28.7	148	Aay74028	Human pro
18	58.5	28.7	148	Aag74543	Human col
19	58.5	28.7	302	Abg03692	Novel hum
20	58.5	28.7	2469	Aae18207	Human MOL
21	58.5	28.7	2469	ADD18192	Human mol
22	58.5	28.7	2471	Aac27065	Human mol
23	58.5	28.7	2471	Aay06816	Human Not
24	58.5	28.7	2471	Aag79774	Human Not
25	58.5	28.7	2471	Abp72572	Human Not

26	58.5	28.7	2471	6	ABR61831	Human Not
27	58.5	28.7	2471	7	ABR61760	Human Not
28	58.5	28.7	2471	7	ADE58245	Human Pro
29	58.5	28.7	2471	7	ADE63707	Human Pro
30	58.5	28.7	2471	7	ADE63703	Human Pro
31	58.5	28.7	2471	7	ADE63715	Human Pro
32	58.5	28.7	2471	7	ADE63711	Human Pro
33	58	28.4	845	2	AAW25717	Mouse gam
34	58	28.4	845	5	ABB57371	Mouse isc
35	57.5	28.2	434	4	ABB66756	Drosophil
36	56	27.5	65	4	AAU4832	Propionib
37	56	27.5	65	6	ABM61351	Propionib
38	55.5	27.2	1140	5	AAM85597	Human imm
39	55.5	27.2	1140	5	AAE27985	Human EGF
40	55.5	27.2	1140	7	ADD18688	Human dis
41	55.5	27.2	1192	7	ADE71305	Novel hum
42	55.5	27.2	2931	4	ABB68229	Drosophil
43	54.5	26.7	78	3	AAQ02842	Human sec
44	54.5	26.7	319	3	AB41889	Human ORF
45	54.5	26.7	580	6	ABJ25525	Aspergill

## ALIGNMENTS

### RESULT 1

AAE02773  
ID AAE02773 standard; peptide; 33 AA.

XX AAE02773;

XX AC AC (first entry)

DT 06-AUG-2001

XX Human Nkp30 receptor intracellular region sequence.

DE Human Nkp30 receptor intracellular region sequence.

XX Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;

KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;

KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;

KW therapy; intracellular region.

XX Homo sapiens.

OS WO200136630-A2.

XX 25-MAY-2001.

XX 15-NOV-2000; 2000WO-EP011697.

XX 15-NOV-1999; 99CA-02288307.

XX 15-NOV-1999; 99US-00440514.

XX (INNA-) INNATE PHARMA SAS.

XX (UYGE-) UNIV GENOVA.

XX Moretta A, Bottino C, Biassoni R;

XX WPI; 2001-329221/34.

XX Novel compound, useful for detection and/or quantifying the presence of

XX NK cells, comprises the amino acid sequences of the Nkp30 molecule.

XX Claim 1; Fig 7B; 83pp; English.

XX The invention relates to human Nkp30 receptor and its corresponding cDNA

XX molecule which is involved in natural cytotoxicity mediated by natural

XX killer (NK) cells and antibodies that identify the same. Nkp30 receptor

XX is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively

XX expressed on the surface of human mature NK cells. Nkp30 and its cDNA are

XX useful for detecting and/or quantifying the presence of NK cells in a

XX biological sample. The invention also provides kits for detecting and/or

XX quantifying the presence of NK cells, for the selective removal of NK

XX cells from a biological sample, for the positive and selective

XX purification of NK cells from a biological sample and for the in vitro











```

PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 30.1%; Score 61.5; DB 3; Length 82;
Best Local Similarity 52.4%; Pred. No. 5.2;
Matches 11; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 3 VYQGRCHCHMGTHCHSSDGP 23
   : ||||| : |||||
Db 48 LPFDGKCH-HLKLHCGSHGP 67

RESULT 6
AAB66269
ID AAB66269 standard; protein; 636 AA.
XX
AC AAB66269;
XX
DT 05-APR-2001 (first entry)
DE Rat TANGO 272 SEQ ID NO: 20.
XX
KW Membrane associated protein; secreted protein; human; mouse; rat;
KW INTERCEPT 340; MANGO 003; MANGO 347; TANGO 272; TANGO 295; TANGO 354;
KW TANGO 378; skeletal disorder; cardiovascular disorder; renal disorder;
KW hematopoietic disorder; neural disorder; hepatic disorder;
KW neoplastic disease.
XX
OS Rattus sp.
XX
PN WO200100673-A1.
XX
PD 04-JAN-2001.
XX
PF 29-JUN-2000; 2000WO-US018198.
XX
PR 30-JUN-1999; 99US-00345464.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Barnes TM, Fraser CC, Wrighton N, Myers P, Busfield SJ, Sharp JD;
DR WPI; 2001-050128/06.
DR N-PSDB; AAF27791.
XX
PT Isolated secreted or transmembrane proteins are used for diagnosis and
PT treatment of neoplastic and hematopoietic disorders e.g. T cell
PT disorders, cancer and tumor.
XX
PS Claim 9; Page 238-240; 294pp; English.
XX
CC The present invention provides the protein and coding sequences for a
CC number of membrane associated and secreted proteins from human, mouse and
CC rat. The proteins are designated INTERCEPT 340, MANGO 003, MANGO 347,

CC TANGO 272, TANGO 295, TANGO 254 and TANGO 378. The proteins are all
CC involved in signal transduction and the sequences can be used in the
CC treatment of cardiovascular, renal, hepatic, neural, neoplastic, skeletal
CC and hematopoietic disorders
XX
SQ Sequence 636 AA;

Query Match 30.1%; Score 61.5; DB 4; Length 636;
Best Local Similarity 44.4%; Pred. No. 39;
Matches 12; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 9 CHCHMGTHCHSSDGPGRVIP---EPRC 32
   ||||| |||||
Db 411 CQCHGATCPQDSCVCVIFGWTGPNC 437

RESULT 7
AAO10665
ID AAO10665 standard; protein; 66 AA.
XX
AC AAO10665;
XX
DT 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO 24557.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
DR WPI; 2001-514838/56.
DR N-PSDB; AAI90596.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 24557; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 66 AA;

Query Match 29.4%; Score 60; DB 4; Length 66;
Best Local Similarity 37.0%; Pred. No. 6.3;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

```

QY 4 YQOKCHCHMGTHCHSSDGRGVIPPEP 30  
 DB 27 HHHHCHCHYHSHHNSQHPPPPPPPP 53

RESULT 8  
 ID AAY06403  
 AC AAY06403 standard; protein; 201 AA.  
 XX AAY06403;  
 XX 20-SEP-1999 (first entry)  
 XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
 KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
 KW signal transduction; immunomodulator; antiinflammatory;  
 KW autoimmune disease; inflammation; gene therapy; diagnosis.  
 XX Homo sapiens.  
 OS  
 XX Key  
 XX Location/Qualifiers  
 FT Peptide  
 FT 1..12  
 FT /note= "leader peptide"  
 FT Protein  
 FT 13..201  
 FT /note= "mature protein"  
 FT Modified-site  
 FT 42  
 FT /note= "N-glycosylated"  
 FT Modified-site  
 FT 68  
 FT /note= "N-glycosylated"  
 FT Modified-site  
 FT 121  
 FT /note= "N-glycosylated"  
 FT Domain  
 FT 139..162  
 FT /note= "transmembrane domain"  
 FT Peptide  
 FT 166..201  
 FT /note= "alternatively spliced C-terminal end"  
 XX WO9923867-A2.  
 XX 20-MAY-1999.  
 XX 05-NOV-1998; 98WO-US023826.  
 XX 07-NOV-1997; 97US-0064761P.  
 XX (BIOJ ) BIOGEN INC.  
 XX Browning J;  
 XX WPI; 1999-418423/35.  
 XX N-PSDB; AAX59349.  
 XX Novel B-cell myelin oligodendrocyte glycoproteins.  
 XX Claim 2; Page 43; 43pp; English.  
 XX This sequence represents human BMOG, a novel member of the B cell myelin  
 XX oligodendrocyte glycoprotein family that is expressed by germinal centre  
 XX B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were  
 XX identified. The protein is present primarily in the spleen, in lymph  
 XX nodes and in germinal centre B cells. It may have immunoregulatory  
 XX functions, and soluble or chimeric fusion proteins of BMOG may be used to  
 XX regulate the immune system in autoimmune or inflammatory disease. Vectors  
 XX comprising BMOG, prokaryotic and eukaryotic host cells, and a method of  
 XX producing BMOG using these transformed host cells are also provided. BMOG  
 XX polypeptides can be used for modulating the immune system of a subject or  
 XX to inhibit signal transduction in a cell expressing BMOG by contacting it  
 XX with a soluble BMOG protein. The nucleic acid can be used for gene  
 XX therapy. The protein can also be used to target a toxin, imaging agent or  
 XX radionuclide to a cell expressing BMOG. (All claimed)  
 XX Sequence 201 AA;

Query Match 29.4%; Score 60; DB 2; Length 201;  
 Best Local Similarity 43.2%; Pred. No. 19;  
 Matches 16; Conservative 2; Mismatches 7; Indels 12; Gaps 2;

QY 1 STVYQKCHCHMGTHCHSSDGR-----GVIPPEPCP 33  
 DB 158 STVYQK-----CLTWKGRRLPAVVPAPLPP 186

RESULT 9  
 ID AAM50359 standard; protein; 1799 AA.  
 AC AAM50359;  
 XX 18-FEB-2002 (first entry) \*  
 XX Mouse laminin-15 beta 2 chain.  
 KW Laminin-15; mouse; retina; eye; therapy; ophthalmological;  
 KW antiinflammatory; rod dystrophy; rod-cone dystrophy;  
 KW retinitis pigmentosa; macular degeneration; retinal detachment.  
 XX Mus musculus.  
 OS WO200183516-A1.  
 XX 08-NOV-2001.  
 XX 01-MAY-2001; 2001WO-US013943.  
 XX 01-MAY-2000; 2000US-0200863P.  
 XX (MASS-) MASSACHUSETTS GEN HOSPITAL.  
 XX Burgeson RE, Brunken W, Champiaud M, Hunter D;  
 XX WPI; 2002-041478/05.  
 XX Novel substantially pure preparation comprising laminin having laminin  
 XX chain alpha 5, beta 2, and gamma 3, useful for treating retinal disorders  
 XX such as retinitis pigmentosa, macular degeneration, retinal detachment.  
 XX Disclosure; Fig 3A; 58pp; English.  
 XX The present sequence is that of the beta 2 chain of mouse laminin-15, a  
 XX novel member of the laminin family that is produced in the retina. The  
 XX retina produces 2 novel laminin trimers: laminin-14 (alpha 4, beta 2,  
 XX gamma 3) and laminin-15 (alpha 5, beta 2, gamma 3). These are expressed  
 XX within the inter-photoreceptor matrix and in the outer plexiform layer,  
 XX and may serve to stabilise retinal synapses. The invention provides  
 XX laminin-15 preparations and cells comprising a nucleic acid encoding the  
 XX laminin alpha 5, beta 2 and gamma 3 chains, and which are capable of  
 XX producing laminin-15. The laminin-15 preparation is used in claimed  
 XX methods of: increasing retina immunophoreceptor matrix stability;  
 XX increasing the stability of retina photoreceptor compounds, especially an  
 XX outer segment, inner segment or synapse; increasing retina adhesion;  
 XX treating a disorder associated with retina degeneration, especially rod  
 XX dystrophy, rod-cone dystrophy, retinitis pigmentosa, macular degeneration  
 XX and retinal detachment; increasing the stability of synapses of the  
 XX central nervous system or peripheral nervous system; stimulating an  
 XX neuroregeneration axon outgrowth or synapse formation; preparing an  
 XX implant, e.g. a catheter, artificial joint, retinal implant, timed  
 XX releasing device, neural cell growth guide or artificial tissue, by  
 XX coating with the laminin-15 preparation; and increasing photosensitivity  
 XX by implanting a tip coated with the laminin-15 preparation into an eye.  
 XX The laminin may be recombinant, and the 3 chains co-expressed in the same  
 XX cell or expressed in different cells  
 XX Sequence 1799 AA;

Query Match 29.4%; Score 60; DB 5; Length 1799;

Best Local Similarity 29.8%; Pred. No. 1.7e+02;  
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

QY 5 YQKCHCHM-----GTHCHSSD-GPRGVPEPRC 32  
: : : : :  
Db 1111 FTGQCHCHAGFGGRTCECOELWGDPLQCRACDCPRG-IDKPOC 1156

RESULT 10  
AAW50895  
ID AAW50895 standard; protein; 1801 AA.  
XX  
AC AAW50895;  
XX  
DT 07-DEC-1998 (first entry)  
XX  
DE Rat laminin B2 chain.  
XX  
KW Laminin; rat; beta-amyloid; amyloidosis; Alzheimer's disease;  
KW Down's syndrome; hereditary cerebral haemorrhage; inflammation;  
KW malignancy; Familial Mediterranean Fever; multiple myeloma;  
KW type II diabetes; prion disease; Creutzfeldt-Jacob disease; CJD;  
KW Gertsmann-Straussler syndrome; kuru; scrapie; haemodialysis;  
KW carpal tunnel syndrome; senile cardiac amyloid polynuropathy;  
KW Familial Amyloidotic Polynuropathy; thyroid carcinoma; diagnosis;  
KW therapy.  
XX  
OS Rattus sp.  
XX  
XX  
XX WO98L5179-A1.  
XX  
PD 16-APR-1998.  
XX  
XX  
XX 08-OCT-1997; 97WO-US018145.  
XX  
XX 08-OCT-1996; 96US-0027981P.  
XX  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Castillo G, Snow AD;  
XX  
XX WPI; 1998-240534/21.  
XX  
XX  
XX Use of laminin and fragments - for developing products for use in the  
XX diagnosis and treatment of amyloid disease, e.g. Alzheimer's disease or  
XX CJD.  
XX  
XX Claim 15; Page 94-97; 132pp; English.

This is the amino acid sequence of the rat laminin B2 chain. The primary  
object of the invention is to use laminin, laminin-derived protein  
fragments and/or laminin-derived polypeptides as potent inhibitors of  
amyloid formation, deposition, accumulation and/or persistence in  
Alzheimer's disease and other amyloidoses. The laminin products (see  
AAW50888-98) may include mammalian laminin A or A1 chain, laminin B1  
or B2 chain, laminin A2 chain (merosin), laminin G1 chain, the globular  
repeats of the laminin A1 chain and the beta-amyloid binding domain of  
the laminin A chain. A claimed method for treating an amyloid disease  
comprises administering a polypeptide having a conformational similarity  
to a fragment of a laminin protein. A method for diagnosing an amyloid  
disease involves determining levels of laminin in a sample. Production of  
laminin or its fourth globular repeat in vivo provides a method for in  
vivo inhibition of beta-amyloid amyloidosis. The products and methods can  
be used for the diagnosis, prognosis, monitoring and treatment of  
amyloidoses such as Alzheimer's disease, Down's syndrome and hereditary  
cerebral haemorrhage with amyloidosis of the Dutch type (where the  
specific amyloid is the beta-amyloid protein), the amyloidosis associated  
with chronic inflammation, various forms of malignancy and Familial  
Mediterranean Fever (AA amyloid or inflammation-association amyloidosis),  
the amyloidosis associated with multiple myeloma and other B-cell  
abnormalities (AL amyloid), the amyloidosis associated with type II  
diabetes (amylin or islet amyloid), the amyloidosis associated with prion  
diseases including Creutzfeldt-Jacob disease, Gertsmann-Straussler

CC syndrome, kuru and animal scrapie (PrP amyloid), the amyloidosis  
CC associated with long-term haemodialysis and carpal tunnel syndrome (beta  
CC 2-microglobulin amyloid), the amyloidosis associated with senile cardiac  
CC amyloid and Familial Amyloidotic Polynuropathy (prealbumin or  
CC transthyretin amyloid), and the amyloidosis associated with endocrine  
CC tumours such as medullary carcinoma of the thyroid (variant of  
XX procalcitonin)

Query Match 29.4%; Score 60; DB 2; Length 1801;  
Best Local Similarity 29.8%; Pred. No. 1.7e+02;  
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

QY 5 YQKCHCHM-----GTHCHSSD-GPRGVPEPRC 32  
: : : : :  
Db 1113 FTGQCHCHAGFGGRTCECOELWGDPLQCRACDCPRG-IDKPOC 1158

RESULT 11  
ADE60383  
ID ADE60383 standard; protein; 1801 AA.  
XX  
AC ADE60383;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein P15800, SEQ ID NO 6292.  
XX  
XX  
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
XX WO2003016475-A2.  
XX  
XX 27-FEB-2003.  
XX  
XX 14-AUG-2002; 2002WO-US025765.  
XX  
XX 14-AUG-2001; 2001US-0312147P.  
XX 01-NOV-2001; 2001US-0346382P.  
XX 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
XX (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX  
XX WPI; 2003-268312/26.  
XX GENBANK; P15800.

New composition comprising two or more isolated polypeptides, useful for  
preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

The invention discloses a composition comprising two or more isolated rat  
or human polynucleotides or a polynucleotide which represents a fragment,  
derivative or allelic variation of the nucleic acid sequence. Also  
claimed are a vector comprising the novel polynucleotide, a host cell  
comprising the vector, a method for identifying a nucleotide sequence  
which is differentially regulated in an animal subjected to pain and a  
kit to perform the method, an array, a method for identifying an agent  
that increases or decreases the expression of the polynucleotide sequence  
subjected to pain, a method for identifying a compound which regulates  
the expression of a polynucleotide sequence which is differentially  
expressed in an animal subjected to pain, a method for identifying a  
compound that regulates the activity of one or more of the  
polynucleotides, a method for producing a pharmaceutical composition, a  
method for identifying a compound or small molecule that regulates the  
activity in an animal of one or more of the polypeptides given in the



Query Match 28.9%; Score 59; DB 3; Length 2813;  
Best Local Similarity 30.6%; Pred. No. 3.4e+02;  
Matches 15; Conservative 4; Mismatches 10; Indels

Qy 4 YQOK-----CHCMG-THCHSDG-----PRGVIPEPRC 32  
Db 710 YDGEIFQPEDIFSDHHTWCYCEDGFHCTTSGGLSLLFNPVLSPPRC 758

Search completed: February 26, 2004, 12:10:15  
Job time : 25.6828 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 11.4668 Seconds  
(without alignments)  
607.670 Million cell updates/sec

Title: US-10-036-444-6  
Perfect score: 204  
Sequence: 1 STVYQKCHCHMGTHCHSSDGRGVPEPRCP 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US05\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	33	US-10-036-444-6	Sequence 6, Appli
2	204	100.0	190	US-10-036-444-2	Sequence 2, Appli
3	61.5	30.1	636	US-09-796-753-100	Sequence 100, Appl
4	61.5	30.1	636	US-09-796-753-124	Sequence 124, Appl
5	60	29.4	1799	US-09-845-593-6	Sequence 6, Appli
6	60	29.4	1801	US-09-938-275-8	Sequence 8, Appli
7	59	28.9	2813	US-09-886-900-2	Sequence 2, Appli
8	58.5	28.7	1418	US-10-106-698-5317	Sequence 5317, Ap
9	58.5	28.7	2469	US-10-190-115-2	Sequence 2, Appli
10	58.5	28.7	2469	US-10-369-072-2	Sequence 2, Appli
11	58	28.4	845	US-09-983-531A-6	Sequence 6, Appli
12	56.5	27.7	969	US-10-052-648A-34	Sequence 34, Appl
13	56.5	27.7	969	US-10-052-648A-35	Sequence 35, Appl
14	55.5	27.2	601	US-10-365-227-20	Sequence 20, Appl
15	55.5	27.2	1140	US-10-092-390-2	Sequence 2, Appli

16	55.5	27.2	1140	15	US-10-052-648A-33	Sequence 33, Appl
17	54.5	26.7	580	14	US-10-128-714-3183	Sequence 3183, Ap
18	54.5	26.7	580	14	US-10-128-714-8183	Sequence 8183, Ap
19	54.5	26.7	1535	14	US-10-188-971-14	Sequence 14, Appl
20	54.5	26.7	1570	14	US-10-189-971-12	Sequence 12, Appl
21	54.5	26.7	1593	14	US-10-189-971-4	Sequence 4, Appl
22	54.5	26.7	1628	14	US-10-189-971-2	Sequence 2, Appli
23	54.5	26.7	2710	13	US-10-153-273-12	Sequence 12, Appl
24	54	26.5	497	10	US-09-796-753-116	Sequence 116, App
25	54	26.5	928	15	US-10-052-648A-4	Sequence 4, Appli
26	54	26.5	928	15	US-10-052-648A-5	Sequence 6, Appli
27	54	26.5	1034	15	US-10-052-648A-31	Sequence 31, Appl
28	54	26.5	1034	15	US-10-052-648A-32	Sequence 32, Appl
29	54	26.5	1037	15	US-10-052-648A-8	Sequence 8, Appli
30	54	26.5	1037	15	US-10-052-648A-10	Sequence 10, Appl
31	54	26.5	1050	10	US-09-796-753-114	Sequence 114, App
32	53.5	26.2	2749	15	US-10-360-101-265	Sequence 265, App
33	53	26.0	597	15	US-10-369-493-5394	Sequence 5394, Ap
34	53	26.0	818	15	US-10-436-715-59	Sequence 59, Appl
35	52.5	25.7	577	13	US-10-072-841-29	Sequence 29, Appl
36	52.5	25.7	615	15	US-10-291-172-192	Sequence 192, App
37	52.5	25.7	965	14	US-10-221-097-51	Sequence 51, Appl
38	52.5	25.7	1374	14	US-10-205-823-93	Sequence 93, Appl
39	52	25.5	165	14	US-10-411-224-101	Sequence 101, App
40	52	25.5	165	15	US-10-047-021-101	Sequence 101, App
41	52	25.5	186	14	US-10-140-184-5	Sequence 5, Appli
42	52	25.5	186	14	US-10-073-333A-5	Sequence 5, Appli
43	52	25.5	255	10	US-09-866-050A-692	Sequence 692, App
44	52	25.5	255	10	US-09-852-472-24	Sequence 24, Appl
45	52	25.5	275	9	US-09-790-264-15	Sequence 15, Appl

## ALIGNMENTS

RESULT 1  
US-10-036-444-6  
; Sequence 6, Application US/10036444  
; Publication No. US20020142445A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; APPLICANT: UNIVERSITA DI GENOVA  
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and  
; TITLE OF INVENTION: antibodies that identify the same"  
; FILE REFERENCE: SEQ-FR-1060  
; CURRENT APPLICATION NUMBER: US/10/036,444  
; PRIOR FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: 09/440,514  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: 09/456,199  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 33  
; TYPE: PRT  
; ORGANISM: Human NK cell  
US-10-036-444-6

Query Watch 100.0%; Score 204; DB 13; Length 33;  
Best Local Similarity 100.0%; Pred. No. 4.5e-17;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHSSDGRGVPEPRCP 33  
Db 1 STVYQKCHCHMGTHCHSSDGRGVPEPRCP 33

RESULT 2  
US-10-036-444-2  
; Sequence 2, Application US/10036444  
; Publication No. US20020142445A1

GENERAL INFORMATION:  
APPLICANT: INNATE PHARMA S.A.S.  
TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same"  
TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same"  
FILE REFERENCE: SEQ-PR-1060  
CURRENT APPLICATION NUMBER: US/10/036,444  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: 09/440,514  
PRIOR FILING DATE: 1999-11-15  
PRIOR APPLICATION NUMBER: 09/456,199  
PRIOR FILING DATE: 1999-12-07  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 2  
LENGTH: 190  
TYPE: PRT  
ORGANISM: Human NK cell  
US-10-036-444-2

Query Match 100.0%; Score 204; DB 13; Length 190;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVYQKCHCHMGTHCHSDGPRGVIPRCP 33  
Db 158 STVYQKCHCHMGTHCHSDGPRGVIPRCP 190

RESULT 3  
US-09-796-753-100  
Sequence 100, Application US/09796753  
Publication No. US20030027998A1

GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
FILE REFERENCE: 7853-227-999  
CURRENT APPLICATION NUMBER: US/09/796,753  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 09/183,175  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 09/223,094  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/223,546  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/224,246  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/259,388  
PRIOR FILING DATE: 1999-02-26  
PRIOR APPLICATION NUMBER: 60/122,458  
PRIOR FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: 09/312,359  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 09/336,536  
PRIOR FILING DATE: 1999-06-18  
PRIOR APPLICATION NUMBER: 09/342,687  
PRIOR FILING DATE: 1999-06-29  
PRIOR APPLICATION NUMBER: 09/345,464  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: 09/365,164  
PRIOR FILING DATE: 1999-07-30  
PRIOR APPLICATION NUMBER: 09/399,723  
PRIOR FILING DATE: 1999-09-20  
PRIOR APPLICATION NUMBER: 09/409,634  
PRIOR FILING DATE: 1999-09-30  
PRIOR APPLICATION NUMBER: 09/471,179  
PRIOR FILING DATE: 1999-12-23  
PRIOR APPLICATION NUMBER: 09/474,071  
PRIOR FILING DATE: 1999-12-29  
PRIOR APPLICATION NUMBER: 09/474,072  
PRIOR FILING DATE: 1999-12-29  
PRIOR APPLICATION NUMBER: 09/514,010

PRIOR FILING DATE: 2000-02-25  
PRIOR APPLICATION NUMBER: 09/516,745  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/572,002  
PRIOR FILING DATE: 2000-05-14  
PRIOR APPLICATION NUMBER: 09/597,993  
PRIOR FILING DATE: 2000-06-19  
PRIOR APPLICATION NUMBER: 09/599,596  
PRIOR FILING DATE: 2000-06-22  
PRIOR APPLICATION NUMBER: 09/630,334  
PRIOR FILING DATE: 2000-07-31  
PRIOR APPLICATION NUMBER: 09/606,565  
PRIOR FILING DATE: 2000-06-29  
PRIOR APPLICATION NUMBER: 09/606,317  
PRIOR FILING DATE: 2000-06-29  
PRIOR APPLICATION NUMBER: 09/665,666  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: 09/677,751  
PRIOR FILING DATE: 2000-09-30  
NUMBER OF SEQ ID NOS: 162  
SEQ ID NO 100  
LENGTH: 636  
TYPE: PRT  
ORGANISM: Rattus sp.  
US-09-796-753-100

Query Match 30.1%; Score 61.5; DB 10; Length 636;  
Best Local Similarity 44.4%; Pred. No. 34;  
Matches 12; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 9 CHCHMGTHCHSDGPRGVIP--EPRC 32  
Db 411 CQCHGATCHPQDQSCVCIEGWTGPNC 437

RESULT 4  
US-09-796-753-124  
Sequence 124, Application US/09796753  
Publication No. US20030027998A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
FILE REFERENCE: 7853-227-999  
CURRENT APPLICATION NUMBER: US/09/796,753  
CURRENT FILING DATE: 2001-03-01  
PRIOR APPLICATION NUMBER: 09/183,175  
PRIOR FILING DATE: 1998-10-30  
PRIOR APPLICATION NUMBER: 09/223,094  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/223,546  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/224,246  
PRIOR FILING DATE: 1998-12-30  
PRIOR APPLICATION NUMBER: 09/259,388  
PRIOR FILING DATE: 1999-02-26  
PRIOR APPLICATION NUMBER: 60/122,458  
PRIOR FILING DATE: 1999-03-01  
PRIOR APPLICATION NUMBER: 09/312,359  
PRIOR FILING DATE: 1999-05-14  
PRIOR APPLICATION NUMBER: 09/336,536  
PRIOR FILING DATE: 1999-06-18  
PRIOR APPLICATION NUMBER: 09/342,687  
PRIOR FILING DATE: 1999-06-29  
PRIOR APPLICATION NUMBER: 09/345,464  
PRIOR FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: 09/365,164  
PRIOR FILING DATE: 1999-07-30  
PRIOR APPLICATION NUMBER: 09/399,723  
PRIOR FILING DATE: 1999-09-20  
PRIOR APPLICATION NUMBER: 09/409,634  
PRIOR FILING DATE: 1999-09-30  
PRIOR APPLICATION NUMBER: 09/471,179  
PRIOR FILING DATE: 1999-12-23

```

; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1998-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1998-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 124
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-796-753-124

```

```

Query Match      30.1%; Score 61.5; DB 10; Length 636;
Best Local Similarity 44.4%; Pred. No. 34;
Matches 12; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

```

```

QY 9 CHCHMGTHCHSSDGRGVIP---EPEC 32
DB 411 CQCHGATCHPQDGSVCIPGTGPNC 437

```

```

RESULT 5
US-09-845-583-6
; Sequence 6, Application US/09845583
; Patent No. US20020142954A1
; GENERAL INFORMATION:
; APPLICANT: Burgeson, Robert
; APPLICANT: Brunken, William Joseph
; APPLICANT: Champlaud, Marie-France
; APPLICANT: Hunter, Dale
; TITLE OF INVENTION: LAMININ 15 AND USES THEREOF
; FILE REFERENCE: 10287-056001
; CURRENT APPLICATION NUMBER: US/09/845,583
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: US 60/200,863
; PRIOR FILING DATE: 2000-05-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1799
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-845-583-6

```

```

Query Match      29.4%; Score 60; DB 9; Length 1799;
Best Local Similarity 29.8%; Pred. No. 1.4e+02;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

```

```

QY 5 YQKCHCHM-----GTHCHSSD-GPRGVPEPRC 32
DB 1111 FTQCHCHAGFGGRTCEQELHWGDPGLQCRACDCDPRG-IDKPCQ 1156

```

```

RESULT 6

```

```

US-09-938-275-8
; Sequence 8, Application US/09938275
; Patent No. US20020111309A1
; GENERAL INFORMATION:
; APPLICANT: Gerardo Caetillo
; APPLICANT: Alan Snow
; TITLE OF INVENTION: Therapeutic and Diagnostic Applications
; TITLE OF INVENTION: of laminin and Laminin-Derived Protein Fragments
; FILE REFERENCE: PROTEO.P03
; CURRENT APPLICATION NUMBER: US/09/938,275
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1801
; TYPE: PRT
; ORGANISM: Rattus No. US20020111309A1vegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Swissprot P15800
; DATABASE ENTRY DATE: 1990-04-01
US-09-938-275-8

```

```

Query Match      29.4%; Score 60; DB 9; Length 1801;
Best Local Similarity 29.8%; Pred. No. 1.4e+02;
Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

```

```

QY 5 YQKCHCHM-----GTHCHSSD-GPRGVPEPRC 32
DB 1113 FTQCHCHAGFGGRTCEQELHWGDPGLQCRACDCDPRG-IDKPCQ 1158

```

```

RESULT 7
US-09-886-900-2
; Sequence 2, Application US/09886900
; Patent No. US20020137051A1
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Wilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48098
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patenlin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/886,900
; FILING DATE: 21-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/896,449
; FILING DATE: 18-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REFERENCE/DOCKET NUMBER: 2115-001226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 248-641-1600
; TELEFAX: 248-641-0270
; TELEX: 287637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2813 amino acids
; TYPE: amino acid

```

```
/
/
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-886-900-2

Query Match      28.9%; Score 59; DB 9; Length 2813;
Best Local Similarity 30.6%; Pred. No. 2.8e+02;
Matches 15; Conservative 4; Mismatches 10; Indels 20; Gaps 3;

Qy 4 YYQK-----CHCHMG-THCHSDG-----PRGVIPPRC 32
Db 710 YDGEIQQPEDIHFDRHHTWCYCEGFWHCTTSGGLGSLLENPNVLSPPRC 758

RESULT 8
US-10-106-698-5317
/ Sequence 5317, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: Patent in Ver. 3.0
/ SEQ ID NO 5317
/ LENGTH: 148
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (18)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5317

Query Match      28.7%; Score 58.5; DB 14; Length 148;
Best Local Similarity 43.5%; Pred. No. 19;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

Qy 10 HCHMGTHCHSDGPRGVIPPRC 32
Db 83 HCHLGYRCHGRPQREGI---PRC 102

RESULT 9
US-10-190-115-2
/ Sequence 2, Application US/10190115
/ Publication No. US2003020739A1
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook, John P. II
/ APPLICANT: Boldog, Ferenc L.
/ APPLICANT: Burgess, Catherine E.
/ APPLICANT: Casman, Stacie J.
/ APPLICANT: Gosse, William M.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Ji, Weizhen
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Mezick, Amanda J.
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Shen, Lei
/ APPLICANT: Shenoy, Suresh G.
/ APPLICANT: Shimkets, Richard A.
/ APPLICANT: Spaderna, Steven K.

Query Match      28.7%; Score 58.5; DB 15; Length 2469;
Best Local Similarity 42.3%; Pred. No. 2.8e+02;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

Qy 7 GKCHCHMGTHC-HSSDGRGVIPPR 31
Db 1345 GQVKRKGECQVHTASGRFCFCPSR 1370

RESULT 10
US-10-369-072-2
/ Sequence 2, Application US/10369072
/ Publication No. US20040014081A1
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zernhusen, Bryan
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Gosse, William M
/ APPLICANT: Sizerkeres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: No. US20040014081A1 Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
```

```

Db          .   647 QGKCHGHGVCSNKNKNCDCD 667

RESULT 12
US-10-052-648A-34
; Sequence 34, Application US/10052648A
; Publication No. US2004000558A1
GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steven
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gunther, Erik
; APPLICANT: Kekuda, Ramesh
; APPLICANT: MacDougall, John R.
; APPLICANT: Mehraban, Foad
; APPLICANT: Patturajan, Meera
; APPLICANT: Rothenberg, Mark
; APPLICANT: Shimkets, Richard
; APPLICANT: Smithson, Glennnda
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Stone, David J.
; APPLICANT: Vernmet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; TITLE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-250 (CURA-550)
; CURRENT APPLICATION NUMBER: US/10/052,648A
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,920
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/284,549
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/303,229
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,605
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/269,098
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 60/264,159
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/271,855
; PRIOR FILING DATE: 2001-02-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 969
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-052-648A-34

Query Match      27.7%; Score 56.5; DB 15; Length 969;
Best Local Similarity 34.4%; Pred. No. 2e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 7; Gaps 1

QY    9 CHCMGWTHCSSDGP-----RGVIEPRCP 33
      ||||| :||| :||| :||| :|||
Db    635 CSCNGASCASABDGACHCTPGWTGLFCTQRCP 666
      ||||| :||| :||| :||| :|||

RESULT 13
US-10-052-648A-35

```

; Sequence 35, Application US/10052648A  
; Publication No. US2004005558A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; APPLICANT: Burgess, Catheline  
; APPLICANT: Casman, Steven  
; APPLICANT: Colman, Steven  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Gunther, Erik  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Mehraban, Fuad  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Zerhusen, Bryan D.  
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
; TITLE OF INVENTION: USING THE SAME  
; FILE REFERENCE: 21402-250 (CURA-550)  
; CURRENT APPLICATION NUMBER: US/10/052,648A  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: 60/262,454  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/272,920  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: 60/284,549  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/303,229  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/262,892  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/263,605  
; PRIOR FILING DATE: 2001-01-23  
; PRIOR APPLICATION NUMBER: 60/269,098  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: 60/264,159  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/271,855  
; PRIOR FILING DATE: 2001-02-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 35  
; LENGTH: 969  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (848)...(889)  
; OTHER INFORMATION: Where Xaa is any amino acid  
US-10-052-648A-35

Query Match 27.7%; Score 56.5; DB 15; Length 969;  
Best Local Similarity 34.4%; Pred. No. 2e+02;  
Matches 11; Conservative 2; Mismatches 12; Indels 7; Gaps 1;

Qy 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33  
Db 635 CSCENGASCSAEDGACHCTPGWTGLFCTQRCP 666

RESULT 14  
US-10-365-227-20  
; Sequence 20, Application US/10365227  
; Publication No. US20030143632A1

; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean A.  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Goodearl, Andrew D.J.  
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER  
; TITLE OF INVENTION: USES  
; FILE REFERENCE: 07334-323001  
; CURRENT APPLICATION NUMBER: US/10/365,227  
; CURRENT FILING DATE: 2003-02-12  
; PRIOR APPLICATION NUMBER: US/09/802,582  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: US 09/128,709  
; PRIOR FILING DATE: 1998-08-04  
; PRIOR APPLICATION NUMBER: US 60/054,645  
; PRIOR FILING DATE: 1997-08-04  
; PRIOR APPLICATION NUMBER: US 09/130,491  
; PRIOR FILING DATE: 1998-08-06  
; PRIOR APPLICATION NUMBER: US 60/054,966  
; PRIOR FILING DATE: 1997-08-06  
; PRIOR APPLICATION NUMBER: US 60/058,108  
; PRIOR FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: US 09/388,280  
; PRIOR FILING DATE: 1999-09-01  
; PRIOR APPLICATION NUMBER: US 09/388,279  
; PRIOR FILING DATE: 1999-09-01  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 601  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-365-227-20

Query Match 27.2%; Score 55.5; DB 14; Length 601;  
Best Local Similarity 34.4%; Pred. No. 1.6e+02;  
Matches 11; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

Qy 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33  
Db 311 CNCHNGAFCSAYDGECKCTPGWTGLYCTQRCP 342

RESULT 15  
US-10-092-390-2  
; Sequence 2, Application US/10092390  
; Publication No. US20030013865A1  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Xuanchuan  
; APPLICANT: Miranda, Maricar  
; TITLE OF INVENTION: No. US20030013865A1 Human EGF-Family Proteins and Polynucleoti  
; FILE REFERENCE: LEX-0317-USA  
; CURRENT APPLICATION NUMBER: US/10/092,390  
; CURRENT FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: US 60/275,013  
; PRIOR FILING DATE: 2001-03-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1140  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-092-390-2

Query Match 27.2%; Score 55.5; DB 14; Length 1140;  
Best Local Similarity 34.4%; Pred. No. 3e+02;  
Matches 11; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

Qy 9 CHCHMGTHCHSSDGP-----RGVPEPRCP 33  
Db 716 CNCHNGAFCSAYDGECKCTPGWTGLYCTQRCP 747

Search completed: February 26, 2004, 12:33:57  
Job time : 13.4668 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 6.38992 Seconds  
(without alignments)  
266.616 Million cell updates/sec

Title: US-10-036-444-6

Perfect score: 204

Sequence: 1 STVYQKCHCHMTHCHSSDGRGVIPRCP 33

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgm2\_6/prodata/2/iaa/5A COMB.pep:\*  
2: /cgm2\_6/prodata/2/iaa/5B COMB.pep:\*  
3: /cgm2\_6/prodata/2/iaa/6A COMB.pep:\*  
4: /cgm2\_6/prodata/2/iaa/6B COMB.pep:\*  
5: /cgm2\_6/prodata/2/iaa/6C COMB.pep:\*  
6: /cgm2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	60	29.4	1799	4	US-09-845-583A-6
2	59	28.9	2813	3	US-08-896-449A-2
3	59	28.9	2813	3	US-09-132-652-2
4	58.5	28.7	1015	1	US-08-537-210A-1
5	58.5	28.7	1015	3	US-09-113-825-1
6	58.5	28.7	2471	1	US-08-185-432-16
7	58.5	28.7	2471	1	US-08-083-590A-19
8	58.5	28.7	2471	3	US-08-532-384-19
9	58.5	28.7	2471	4	US-08-899-232-1
10	54.5	26.7	2710	2	US-08-568-459A-12
11	54.5	26.7	2710	2	US-08-487-826B-12
12	54.5	26.7	2710	4	US-09-210-288-12
13	54.5	26.7	3060	2	US-08-487-826B-14
14	52.5	25.7	577	2	US-07-728-215-29
15	52.5	25.7	577	4	US-08-938-085A-29
16	52.5	25.7	577	4	US-10-072-844-29
17	52.5	25.7	577	4	US-10-072-838-29
18	52.5	25.7	577	4	US-10-072-841A-29
19	52.5	25.7	799	1	US-08-054-077C-2
20	52	25.5	278	4	US-09-724-864-52
21	52	25.5	420	4	US-09-252-991A-27434
22	52	25.5	470	4	US-09-252-991A-27804
23	51	25.0	70	4	US-09-381-546-21
24	51	25.0	207	3	US-08-974-023-47
25	51	25.0	207	3	US-08-795-445A-47
26	51	25.0	207	3	US-08-795-447A-47
27	51	25.0	207	3	US-08-974-186-47

Sequence 47, Appl  
Sequence 133, App  
Sequence 47, Appl  
Sequence 2, Appl  
Sequence 9, Appl  
Sequence 9, Appl  
Sequence 9, Appl  
Sequence 24401, A  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 8, Appl  
Sequence 11, Appl  
Sequence 192, App  
Sequence 38, Appl  
Sequence 38, Appl  
Sequence 325, App

28 51 25.0 207 3 US-08-795-446B-47  
29 51 25.0 207 4 US-08-706-945D-133  
30 51 25.0 207 4 US-08-577-788C-47  
31 51 25.0 325 1 US-08-292-549-2  
32 51 25.0 325 3 US-09-042-785A-9  
33 51 25.0 325 5 PCT-US91-02207-2  
34 50.5 24.8 298 3 US-09-080-044-9  
35 50.5 24.8 298 4 US-09-531-857A-9  
36 50.5 24.8 771 4 US-09-252-991A-24401  
37 50 24.5 277 1 US-08-690-457-5  
38 50 24.5 277 2 US-08-628-187-5  
39 50 24.5 277 3 US-08-432-071-3  
40 50 24.5 1798 4 US-09-845-583A-8  
41 50 24.5 1798 4 US-09-561-709B-11  
42 49.5 24.3 299 3 US-09-188-930-192  
43 49 24.0 115 3 US-08-513-974B-38  
44 49 24.0 115 4 US-09-461-436B-38  
45 49 24.0 132 3 US-08-513-974B-325

#### ALIGNMENTS

##### RESULT 1

US-09-845-583A-6  
; Sequence 6, Application US/09845583A  
; Patent No. 6635616

; GENERAL INFORMATION:

; APPLICANT: Burgeson, Robert

; APPLICANT: Brunken, William Joseph

; APPLICANT: Champlaud, Marie-France

; APPLICANT: Hunter, Dale

; TITLE OF INVENTION: LAMININ 15 AND USES THEREOF

; FILE REFERENCE: 10287-056001

; CURRENT APPLICATION NUMBER: US/09/845.583A

; PRIOR FILING DATE: 2001-04-30

; PRIOR FILING DATE: 2000-05-01

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 1799

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-845-583A-6

Query Match 29.4%; Score 60; DB 4; Length 1799;

Best Local Similarity 29.8%; Pred. No. 25;

Matches 14; Conservative 5; Mismatches 8; Indels 20; Gaps 3;

Cy 5 YQKCHCHM-----GTHCHSSD-GPRGVIPRCP 32

Db 1111 FTGQCHAGFGGRTCTCQELYWGDPGLQCRACDCPRG-IDKQPC 1156

##### RESULT 2

US-08-896-449A-2

; Sequence 2, Application US/08896449A

; Patent No. 6040143

; GENERAL INFORMATION:

; APPLICANT: Ventr, Patrick J

; APPLICANT: Yuzbasiyan-Gurkan, Vilma

; APPLICANT: Schall, William D

; APPLICANT: Brewer, George J

; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND

; NUMBER OF SEQUENCE: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

; STREET: 5445 Corporate Drive

; CITY: Troy

; STATE: Michigan

; COUNTRY: USA





```

; ; COMPUTER: IBM Compatible
; ; OPERATING SYSTEM: DOS
; ; SOFTWARE: FastSeq Version 2.0
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/09/113,825
; ; FILING DATE:
; ; CLASSIFICATION:
; ; PRIOR APPLICATION DATA:
; ; APPLICATION NUMBER: 08/537,210
; ; FILING DATE: 29-SEP-1995
; ; ATTORNEY/AGENT INFORMATION:
; ; NAME: Misrock, S. Leslie
; ; REGISTRATION NUMBER: 18,872
; ; REFERENCE/DOCKET NUMBER: 7326-027
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: 212-790-9090
; ; TELEFAX: 212-869-8864
; ; TELEX: 66141 PENNIE
; ; INFORMATION FOR SEQ ID NO: 1:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 1015 amino acids
; ; TYPE: amino acid
; ; STRANDEDNESS:
; ; TOPOLOGY: unknown
; ; MOLECULE TYPE: protein
; ; FEATURE:
; ; NAME/KEY: hum N (Human No. 6149902ch 2)
; ; LOCATION: 1155...2169
; ; OTHER INFORMATION: Highly conserved ankyrin repeat
; ; OTHER INFORMATION: region of No. 6149902ch
US-09-113-825-1

Query Match 28.7%; Score 58.5; DB 3; Length 1015;
Best Local Similarity 42.3%; Pred. No. 21;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVPEPR 31
Db 193 GQVKCRKGECVHTASGPRCFCPSR 218

RESULT 6
US-08-185-432-16
; Sequence 16, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Atavanis-Teakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Jian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: E30
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872

```

```

; ; REFERENCE/DOCKET NUMBER: 7326-006
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (212) 790-9090
; ; TELEFAX: (212) 869-8864/9741
; ; TELEX: 66141 PENNIE
; ; INFORMATION FOR SEQ ID NO: 16:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 2471 amino acids
; ; TYPE: amino acid
; ; TOPOLOGY: unknown
; ; MOLECULE TYPE: protein
US-08-185-432-16

Query Match 28.7%; Score 58.5; DB 1; Length 2471;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVPEPR 31
Db 1347 GQVKCRKGECVHTASGPRCFCPSR 1372

RESULT 7
US-08-083-590A-19
; Sequence 19, Application US/08083590A
; Patent No. 5786158
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Teakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,590A
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-083-590A-19

Query Match 28.7%; Score 58.5; DB 1; Length 2471;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 3; Mismatches 11; Indels 1; Gaps 1;

QY 7 GKCHCHMGTHC-HSSDGRGVPEPR 31
Db 1347 GQVKCRKGECVHTASGPRCFCPSR 1372

```

```

; SEQ ID NO 1
; LENGTH: 2471
; TYPE: PRT
; ORGANISM: Homo sapiens

```

RESULT 11  
US-08-487-826B-12  
; Sequence 12, Application US/08487826B  
; Patent No. 5993827

GENERAL INFORMATION:  
APPLICANT: Sim, Kim L.  
APPLICANT: Chitnis, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhaun  
APPLICANT: Wellens, Thomas E.  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CPI  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE: Plasmodium falciparum  
US-08-487-826B-12

Query Match 26.7%; Score 54.5; DB 2; Length 2710;  
Best Local Similarity 40.0%; Pred. No. 1.8e+02;  
Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 5 YQKCKCHMGTHCHS-SDGPRGVIPRCP 33  
DB 1700 YPEKDCYQGHVPSIPPPPPVQPOPEAP 1729

RESULT 12  
US-09-210-288-12  
Sequence 12, Application US/09210288  
Patent No. 632026  
GENERAL INFORMATION:  
APPLICANT: Sim, Kim L.  
APPLICANT: Chitnis, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhaun  
APPLICANT: Wellens, Thomas E.  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California

COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,288  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fuller, Michael  
REGISTRATION NUMBER: 36,516  
REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2710 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE: Plasmodium falciparum  
US-09-210-288-12

Query Match 26.7%; Score 54.5; DB 4; Length 2710;  
Best Local Similarity 40.0%; Pred. No. 1.8e+02;  
Matches 12; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

QY 5 YQKCKCHMGTHCHS-SDGPRGVIPRCP 33  
DB 1700 YPEKDCYQGHVPSIPPPPPVQPOPEAP 1729

RESULT 13  
US-08-487-826B-14  
Sequence 14, Application US/08487826B  
Patent No. 5993827  
GENERAL INFORMATION:  
APPLICANT: Sim, Kim L.  
APPLICANT: Chitnis, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhaun  
APPLICANT: Wellens, Thomas E.  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CPI

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3060 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-826B-14

```

Query Match 26.7%; Score 54.5; DB 2; Length 3060;  
Best Local Similarity 40.0%; Pred. No. 2.1e+02;  
Matches 12; Conservative 2; Mismatches 15; Indels 1

## RESULT 14

US-07-728-215-29 ; Sequence 29, Application US/07728215 ; Patent No. 5962643

Query Match	25.7%	Score 52.5;	DB 2;	Length 577;
Best Local Similarity	34.2%	Pred. No. 68;		
Matches 13;	Conservative	2;	Mismatches	14;
				Indels

RESULT 15

US-08-938-085A-29  
; Sequence 29, Application US/08938085A  
; Patent No. 6339148  
; GENERAL INFORMATION:

Query Match	25.7%;	Score 52.5;	DB 4;	Length 577;
Best Local Similarity	34.2%;	Pred. No. 68;		

Search completed: February 26, 2004, 12:16:18  
Job time : 7.38992 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:05:08 ; Search time 2.2679 Seconds  
(without alignments)  
636.214 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPPIRLTLEGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: Piri:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46.5	54.7	155	2 F71637	peptidoglycan-asso
2	46.5	54.7	155	2 H97849	peptidoglycan-asso
3	45	52.9	353	2 T01256	probable anthocyan
4	43	50.6	305	2 T36261	hypothetical prote
5	43	50.6	514	2 A49838	site-specific reco
6	43	50.6	514	2 AF1988	fdxN element site-
7	43	50.6	687	2 AD2347	ribonuclease E lim
8	42	49.4	101	2 G75512	conserved hypothet
9	42	49.4	372	1 S28296	hypothetical prote
10	42	49.4	478	1 S60754	transcription fact
11	42	49.4	486	1 TVHUEG	transforming prote
12	42	49.4	535	2 S75159	hypothetical prote
13	42	49.4	1317	2 F83310	conserved hypothet
14	41.5	48.8	60	2 S63769	short neurotoxin -
15	41	48.2	272	2 S36641	flagellar motor ap
16	41	48.2	347	2 S71436	farnesyl-pyrophosp
17	41	48.2	368	2 S67507	morphogen lefty pr
18	41	48.2	517	2 C98840	conserved hypothet
19	41	48.2	597	2 T08681	adenosylhomocystei
20	41	48.2	788	2 AG0786	secreted effector
21	41	48.2	805	2 T24399	hypothetical prote
22	40	47.1	196	2 H64968	acetyl CoA acetyl
23	40	47.1	209	2 B93329	probable acyltrans
24	40	47.1	316	2 E75566	probable repressor
25	40	47.1	341	2 S42699	hypothetical prote
26	40	47.1	341	2 B35769	cAMP response elem
27	40	47.1	341	2 S03343	CAMP response elem
28	40	47.1	349	2 D75419	proline dipeptidas
29	40	47.1	396	2 E64049	cystathionine beta

30	40	47.1	443	2 H96841	hypothetical prote
31	40	47.1	664	2 T48258	kinesin-like prote
32	40	47.1	690	2 C69208	DNA helicase relat
33	40	47.1	3707	2 S18262	heparan sulfate pr
34	39.5	46.5	1694	2 S50865	sialoadhesin - mou
35	39	45.9	107	2 F81747	hypothetical prote
36	39	45.9	108	2 T13133	hypothetical prote
37	39	45.9	293	2 S64925	protein gp46 - pha
38	39	45.9	371	2 AH1263	hypothetical prote
39	39	45.9	434	2 F41880	probable tRNA (5-m
40	39	45.9	515	2 E71447	47.87K diazotroph
41	39	45.9	527	2 S46888	probable P8II D1 p
42	39	45.9	916	2 G75417	hypothetical prote
43	39	45.9	1197	2 A91035	SNF2/Rad54 helicase
44	39	45.9	1197	2 C85879	probable sensor fo
45	38.5	45.3	185	2 AB2241	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

F71637

peptidoglycan-associated lipoprotein precursor (pal) RP771 - Rickettsia prowazekii  
C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000

C:Accession: F71637 Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark,

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499; PMID:9823893

A:Accession: F71637

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-155 <AND>

A:Cross-references: GB:AJ235273; GB:AJ235269; NID:G3861237; PIDN:CAAL5198.1; PID:G38612

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: pal; RP771

C:Superfamily: outer membrane protein A

Query Match 54.7%; Score 46.5; DB 2; Length 155;

Best Local Similarity 50.0%; Pred. No. 2.5;

Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy 1 WVSQPPPIRLTLEGSC 15

Db 74 WLSKHPVKVTVEGHC 89

##### RESULT 2

H97849

peptidoglycan-associated lipoprotein precursor [imported] - Rickettsia conorii (strain

C:Species: Rickettsia conorii

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 03-Jun-2002

C:Accession: H97849

R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R

Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:Reference number: A97700; MUID:21442074; PMID:11557893

A:Accession: H97849

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <KOR>

A:Cross-references: GB:AE006914; PIDN:AAL03738.1; PID:G15620330; GSPDB:GN00173

C:Genetics:

A:Gene: pal

C:Superfamily: outer membrane protein A

Query Match 54.7%; Score 46.5; DB 2; Length 155;

Best Local Similarity 50.0%; Pred. No. 2.5;

Matches 8; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy            1 WVSQPPEIR-TLEGSC 15  
             |::|::|::|  
Db            74 WLSKHPEVKVTVEGHC 89

### RESULT 3

T01256

Probable anthocyanidin synthase [imported] - Arabidopsis thaliana  
N;Alternate names: SGR1 protein homolog F16M14.17  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001  
C/Accession: T01256; F84802  
R/Rounseley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, A.  
Submitted to the EMBL Data Library, July 1998  
A/Description: Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence.  
A/Reference number: Z14213  
A/Accession: T01256  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-353 <ROI>  
A/Cross-references: EMBL:AC003028; NID:G3335356; PID:G3353572  
A/Experimental source: cultivar:Columbia  
R/Lin, X.; Kaul S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Varaken, S.E.; Umayam, L.; Tallon, L.;  
euser, D.; Nierman, W.C.; White, C.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A/Reference number: A84420; MUID:20083487; PMID:10617197

## A;Status: preliminar

A:Molecule type: DNA  
A:Residues: 1-353 <STO>  
A:Cross-references: GB:AE002093; NID:g3335372; PIDN:AAC27173.1; GSPDB:GNO0139  
C:GeneIds:  
A:Gene: F16M14.17; At2g38240  
A:Map position: 2  
A:Introns: 160/2; 271/3  
C:Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

```
Query Match          52.9%; Score 45; DB 2; Length 353;
Best Local Similarity 72.7%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```

QY      1 WVSQPPEIRTL 11
      152 WPSOPPKIREL 16
Dp

```

#### RESULT 4

T36261  
 T36261  
 hypothetical protein SCE68.11 - Streptomyces coelicolor  
 C;Species: Streptomyces coelicolor  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C;Accession: T36261  
 R;Murphy, L.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, June 1999  
 A;Reference number: Z21576  
 A;Accession: T36261  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-305 <MUR>  
 A;Cross-references: EMBL:AL079345; PIDN:CAB45347.1; GSPDB:GN00070; SCOEDB:SCE68.11  
 A;Experimental source: strain A3(2)  
 C;Genetics:  
 A;Gene: SCOEDB:SCE68.11

```
Query Match          50.6%; Score 43; DB 2; Length 305;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7: Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

QY	1	WVSQPPEIRTL	11
		: :	
Db	140	WVGRRPREVRTL	15

## RESULT 5

A49838  
 site-specific recombinase XisF - Anabaena sp. (strain PCC 7120)  
 C:Species: Anabaena sp.  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
 C:Accession: A49838  
 R:Carasco, C.D.; Ramaswamy, K.S.; Ramasubramanian, T.S.; Golden, J.W.  
 Genes Dev. 8, 74-83, 1994  
 A:Title: Anabaena xisF gene encodes a developmentally regulated site-specific recombinase  
 A:Reference number: A49838; MUID:94116860; PMID:8288129  
 A:Accession: A49838  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-514 <CAR>  
 A:Cross-references: GB:L23220; NID:G349678; PIDN:AAA16762.1; PID:G349679  
 C:Genetics:  
 A:Gene: xisF

Query Match	50.6%	Score 43;	DB 2;	Length 514;
Best Local Similarity	53.8%;	Pred. No. 32;		
Matches	7;	Conservative	3;	Mismatches 3;
				Indels 0;
				Gaps 0;

2 VSOPPEIRTLGS 14

414 VEEPEVKTI.BAS 426

## RESULT 6

AF1988  
 fixN element site-specific recombinase [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: AF1988  
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch  
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata,  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AF1988  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-514 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BAE73414.1; PID:gl7130804; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: xisF

## Query Match

Query Match 50.6%; Score 45; DB 2; Length 314;  
Best Local Similarity 53.8%; Pred. No. 32;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

2 VSOPPEIRTLGS 14

Db 414 VEEPPVKTLLRAS 426

## RESULT 7

AD2347  
AD2347 ribonuclease E [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C;Accession: AD2347  
R;Kanaketo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch,  
Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
A;Reference number: AB1807; PMID:21595285; PMID:11759840  
A;Accession: AD2347  
A;Status: preliminary

Qy 2 VSQPPEIRTEGS 14  
||| : |||  
Db 21 VSAPPEYKRLDGS 33

A/Map position: 21q22.2-21q22.2  
 B/Crosses/References: Q9B-113084; Q91N1A; Q93060  
 C/Superfamily: transcription factor; ets DNA-binding domain homology; ets RII regul  
 D/Keywords: alternative splicing; DNA binding; nucleus; proto-oncogene; transcription f  
 E/F1-231,256-486/Product: transforming protein erg-2 #status predicted <EG2>  
 F/100-231,256-486/Product: transforming protein erg-1 #status predicted <EG2>  
 F/126-200/Domain: ets RII regulatory region homology  
 F/126-200/Domain: ets RII regulatory region homology <ETR>



F:320-398/Domain: ets DNA-binding domain homology <ETS>

Query Match 49.4%; Score 42; DB 1; Length 486;  
Best Local Similarity 46.7%; Pred. No. 44;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 WVSQPPPIRTLEGGC 15  
|:|||||:|:  
Db 63 WLSQPPPARVTIMKEC 77

RESULT 12

S75159

hypothetical protein slr1771 - Synecchocystis sp. (strain PCC 6803)

C:Species: Synecchocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75159

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-535 <KAN>

A:Cross-references: EMBL:D90903; GB:AB001339; NID:G1652127; PIDN:BAAL1073.1; PID:d101780

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 49.4%; Score 42; DB 2; Length 535;  
Best Local Similarity 85.7%; Pred. No. 48;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPE 7  
|:|||||:  
Db 203 WISQPE 209

RESULT 13

F83310

conserved hypothetical protein PA2694 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: F83310

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: F83310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1317 <STO>

A:Cross-references: GB:AE004696; GB:AE004091; NID:G9948750; PIDN:AAG06072.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2694

C:Superfamily: rhes protein

Query Match 49.4%; Score 42; DB 2; Length 1317;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WVSQPPPIRTLEGS 14  
|:|||||:  
Db 155 WAAVPAELQTEGS 168

RESULT 14

S68769

short neurotoxin - black-banded coral snake

N:Alternate names: alpha-neurotoxin

C:Species: Micurus nigrocinctus

C:Date: 25-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 18-Jul-2001

C:Accession: S68769

R:Rosso, J.P.; Vargas-Rosso, O.; Gutierrez, J.M.; Rochat, H.; Bougis, P.E.

Eur. J. Biochem. 238, 231-239, 1996

A:Title: Characterization of alpha-neurotoxin and phospholipase A(2) activities from Mi

in from Micurus nigrocinctus nigrocinctus.

A:Reference number: S68769; MUID:96248443; PMID:8665942

A:Accession: S68769

A:Molecule type: protein

A:Residues: 1-60 <ROS>

A:Experimental source: Micurus nigrocinctus (black-banded coral snake)

C:Superfamily: snake toxin

C:Keywords: neurotoxin; venom

F:3-22,17-39,41-52,53-58/Disulfide bonds: #status predicted

Query Match 48.8%; Score 41.5; DB 2; Length 60;  
Best Local Similarity 64.3%; Pred. No. 6.3;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 3 SQPPPIRTL-EGSC 15  
|:|||||:  
Db 9 SQPPPIKTCSEGC 22

RESULT 15

S39641

flagellar motor apparatus homolog ytxD - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 08-Jun-1994 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000

C:Accession: S39641; B70003

R:Grundy, F.J.; Waters, D.A.; Takova, T.Y.; Henkin, T.M.

Mol. Microbiol. 10, 259-271, 1993

A:Title: Identification of genes involved in utilization of acetate and acetoin in Baci

A:Reference number: S39641; MUID:95020526; PMID:7934817

A:Accession: S39641

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-272 <GRU>

A:Cross-references: GB:117309; NID:9861173; PIDN:AA668282.1; PID:9348048

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berle

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

teck, J.; Harwood, C.R.; Hanaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

Y, M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, A.; Liu, H.; Magada, S.; Mauee

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:198044033; PMID:9384377

A:Accession: B70003

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-272 <KUN>

A:Cross-references: GB:299119; GB:AL009126; NID:G2635411; PIDN:CAB14951.1; PID:G2635457

A:Experimental source: strain 168

C:Genetics:

A:Gene: ytxD

C:Superfamily: flagellar motor rotation protein

C:Keywords: transmembrane protein

Query Match 48.2%; Score 41; DB 2; Length 272;  
Best Local Similarity 42.9%; Pred. No. 36;  
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 WVSOPPEIRTEGS 14  
Db 255 WQOPKQVTKKGS 268

Search completed: February 26, 2004, 12:14:59  
Job time : 5.2679 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:12 ; Search time 1.27321 Seconds  
(without alignments)  
613.452 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPBIRTEGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681.

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	51.8	688	1 EOMD_MOUSE	O54839 mus musculus
2	43	50.6	692	1 EOMD_XENLA	F79944 xenopus lae
3	42	49.4	372	1 YLFI_CABEL	Q03571 caenorhabdi
4	42	49.4	383	1 TRMU_PASMU	Q96143 pasteurilla
5	42	49.4	462	1 ERG_HUMAN	P11308 homo sapien
6	42	49.4	478	1 ERG_CHICK	Q90837 gallus gall
7	41.5	48.8	60	1 NX5I_MICNI	P80548 micrurus ni
8	41	48.2	272	1 YTXD_BACSU	P39063 bacillus su
9	41	48.2	347	1 FPPS_NEUCR	Q92250 neurospora
10	41	48.2	368	1 LFTB_MOUSE	P57785 mus musculus
11	41	48.2	368	1 TGF4_MOUSE	Q64280 mus musculus
12	41	48.2	686	1 EOMD_HUMAN	O95936 homo sapien
13	41	48.2	1411	1 Y297_HUMAN	O15040 homo sapien
14	40	47.1	196	1 WBBJ_ECOLI	P37750 escherichia
15	40	47.1	341	1 CREB_HUMAN	P16220 homo sapien
16	40	47.1	341	1 CREB_MOUSE	Q01147 mus musculus
17	40	47.1	341	1 CREB_RAT	P15337 rattus norv
18	40	47.1	396	1 METC_HAEIN	Q45227 haemophilus
19	40	47.1	690	1 HELS_METTH	O26901 methanobact
20	40	47.1	3707	1 PGBM_MOUSE	Q05793 mus musculus
21	39.5	46.5	1694	1 SN_MOUSE	Q62230 mus musculus
22	39.5	46.5	1709	1 SN_HUMAN	Q95222 homo sapien
23	39	45.9	371	1 TRMU_LISMO	Q8Y114 listeria mo
24	39	45.9	434	1 RP54_RHOSH	Q01194 rhodobacter
25	39	45.9	527	1 YB64_YEAST	P38314 saccharomyc
26	39	45.9	942	1 M3KE_MOUSE	Q9WU16 mus musculus
27	39	45.9	1017	1 TRL3_HUMAN	Q9HCF6 homo sapien
28	39	45.9	1197	1 EVGS_ECO57	P58402 escherichia
29	38	44.7	187	1 RK24_TCBAC	Q02764 nicotiana t
30	38	44.7	334	1 GTR8_BOVIN	P58354 bos taurus
31	38	44.7	369	1 PROB_NEIMA	Q9JUK7 neisseria m
32	38	44.7	369	1 PROB_NEIME	Q9J292 neisseria m
33	38	44.7	409	1 LISI_BOVIN	P43033 bos taurus

RESULT 1

ID	EOMD_MOUSE	STANDARD;	PRT;	688 AA.
AC	O54839; Q9QYQ7;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Eomesodermin homolog.			
GN	EOMES OR TBR2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=99337662; PubMed=10407135;			
RA	Kimura N., Nakashima K., Ueno M., Taga T.;			
RT	"A novel mammalian T-box-containing gene, Tbr2, expressed in mouse developing brain."			
RL	Brain Res. Dev. Brain Res. 115:183-193 (1999).			
RN	[2]			
RP	SEQUENCE OF 278-457 FROM N.A.			
RX	MEDLINE=98163742; PubMed=9503012;			
RA	Wattler S., Russ A., Evans M., Nehls M.;			
RT	"A combined analysis of genomic and primary protein structure defines the phylogenetic relationship of new members of the T-box family.";			
RL	Genomics 48:24-33 (1998).			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Potential).			
CC	-1- SIMILARITY: Contains 1 T-box domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; AB031037; BAA83416.1; -			
DR	EMBL; AF013281; AAC16233.1; -			
DR	HSPB; P24781; 1XBR.			
DR	MGD; MGI:1201683; Eomes.			
DR	InterPro; IPR008967; P53-like.			
DR	Profam; PF00907; T-box; 1.			
DR	PRINTS; PR00937; TBOX.			
DR	SMART; SM00425; TBOX; 1.			
DR	PROSITE; PS01283; TBOX 1; 1.			
DR	PROSITE; PS01264; TBOX 2; 1.			
DR	PROSITE; PS02522; TBOX 3; 1.			
KW	Developmental protein; Transcription regulation; DNA-binding;			
KW	Nuclear protein.			
FT	DOMAIN 27 41			POLY-GLY.
FT	DNA BIND 278 458			T-BOX.
FT	DOMAIN 383 386			POLY-ASN.

## ALIGNMENTS

P43034 homo sapien  
P43035 mus musculus  
O16999 caenorhabdi  
Q01603 drosophila  
Q99558 homo sapien  
P17139 caenorhabdi  
Q14674 homo sapien  
O88799 mus musculus  
P20956 xenopus lae  
Q9YF53 aeropyrum p  
P01422 naja naja a  
P25675 naja naja h

34 38 44.7 409 1 LISI\_HUMAN  
35 38 44.7 409 1 LISI\_MOUSE  
36 38 44.7 458 1 ZPRI\_CABEL  
37 38 44.7 690 1 FERO\_DROME  
38 38 44.7 947 1 M3KE\_HUMAN  
39 38 44.7 1758 1 CAL14\_CABEL  
40 38 44.7 1795 1 ESPI\_HUMAN  
41 38 44.7 5376 1 ZAN\_MOUSE  
42 37.5 44.1 136 1 HVOI\_XENLA  
43 37.5 44.1 155 1 RL19\_AERPE  
44 37 43.5 61 1 NXS2\_NAJHA  
45 37 43.5 61 1 NXS2\_NAJHH

```
SQ SEQUENCE 688 AA; 72638 MW; 197B0B99E920B82B CRC64;
Query Match 51.8%; Score 44; DB 1; Length 688;
Best Local Similarity 42.9%; Pred. No. 8.7;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 WVSQPPPEIRTLGGS 14
   ||| ||| :|
Db 620 WNETPPSIKSLDSS 633

RESULT 2
EOMD XENLA
ID EOMD XENLA STANDARD; PRT; 692 AA.
AC P79944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Eomesodermin.
GN EOMES.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gastrula;
RX MEDLINE=97133207; PubMed=8978604;
RT Ryan K., Garrett N., Mitchell A., Gurdon J.B.;
RL Cell 87:989-1000(1996).
CC -|- FUNCTION: Involved in mesoderm differentiation. Activates WNT8,
CC brachyury, CHD and MIX.1 expression.
CC -|- SUBCELLULAR LOCATION: Nuclear (potential).
CC -|- DEVELOPMENTAL STAGE: First expressed at or just after midblastula
CC transition (stage 8). Maximally expressed at stage 10 as an
CC equatorial mesoderm band, more prominently on the dorsal side
CC and around the invaginating dorsal lip.
CC -|- INDUCTION: By activin.
CC -|- DOMAIN: CONTAINS 13 S-P-X-X REPEATS.
CC -|- SIMILARITY: Contains 1 T-box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U75996; AAC60061.1; -.
CC HSP; P24781; 1XER.
CC InterPro; IPR008967; P53-like.
CC InterPro; IPR001699; TF_T-box.
CC Pfam; PF00907; T-box; 1.
CC PRINTS; PR00937; TBOX.
CC SMART; SM00425; TBOX; 1.
CC PROSITE; PS01283; TBOX_1; 1.
CC PROSITE; PS01284; TBOX_2; 1.
CC PROSITE; PS02522; TBOX_3; 1.
CC Developmental protein; Transcription regulation; DNA-binding;
CC Nuclear protein; Repeat.
CC DNA BIND 263 443
FT T-BOX.
SQ SEQUENCE 692 AA; 75943 MW; 9D129A67F6357989 CRC64;
Query Match 50.6%; Score 43; DB 1; Length 692;
Best Local Similarity 34.8%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 3; Indels 8; Gaps 1;

QY 1 WVSQPPPEIRTLGGS-----GSC 15
   ||| ||| :|
Db 626 WNETPPSIKSLDSSGVYTGAC 648

SQ SEQUENCE 688 AA; 72638 MW; 197B0B99E920B82B CRC64;
Query Match 51.8%; Score 44; DB 1; Length 688;
Best Local Similarity 42.9%; Pred. No. 8.7;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 WVSQPPPEIRTLGGS 14
   ||| ||| :|
Db 620 WNETPPSIKSLDSS 633

RESULT 3
YLF1 CAEL
ID YLF1 CAEL STANDARD; PRT; 372 AA.
AC Q03571;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical RNA-binding protein C40H1.1 in chromosome III.
GN C40H1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Jier M.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -|- SIMILARITY: Contains 2 RNA recognition motif (RRM) domains.
CC -|- SIMILARITY: TO DROSOPHILA OVARIAN PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z19154; CAA79552.1; -.
CC PIR; S28296; S28296.
CC WormPep; C40H1.1; CE00109.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 2.
CC SMART; SM00360; RRM; 2.
CC PROSITE; PS00102; RRM; 2.
CC PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
CC Hypothetical protein; RNA-binding; Repeat.
CC FT DOMAIN 58 161
CC RNA-BINDING (RRM) 1.
CC DOMAIN 178 249
CC RNA-BINDING (RRM) 2.
CC SEQUENCE 372 AA; 42393 MW; 8016734389D97E18 CRC64;
Query Match 49.4%; Score 42; DB 1; Length 372;
Best Local Similarity 61.5%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 VSQPPPEIRTLGGS 14
   ||| ||| :|
Db 21 VSAPPEVKRLDGS 33

RESULT 4
TRMU PASMU
ID TRMU PASMU STANDARD; PRT; 383 AA.
AC Q9C1A3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
```



Query Match 49.4%; Score 42; DB 1; Length 462;  
Best Local Similarity 46.7%; Pred. No. 13;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WVSQPPERTLEGSC 15  
DB 63 WLSQPPARVIRMEC 77

## RESULT 6

ERG\_CHICK ID\_ERG\_CHICK STANDARD; PRT; 478 AA.  
AC Q30837;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Transcriptional regulator Erg.  
GN ERG.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archozoa; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RX MEDLINE=95329425; PubMed=7605748;  
RA Dordain P., Dewitte F., Debliens X., Stehelin D.,  
RA Dutreque-Coquillaud M.;  
RT "Mesodermal expression of the chicken erg gene associated with  
RT precartilaginous condensation and cartilage differentiation.";  
RL Mech. Dev. 50:17-28(1995).  
CC -1- FUNCTION: Acts as a transcriptional activator.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MESODERM- AND, TO A LESSER  
CC EXTENT, IN ECTODERM-DERIVED TISSUES.  
CC -1- SIMILARITY: Belongs to the ETS family.  
CC -1- SIMILARITY: Contains 1 pointed (PNT) domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [licenses@isb-sib.ch](mailto:licenses@isb-sib.ch)).  
CC  
CC EMBL; X77159; CAA54404.1; --  
CC PIR; S60754; S60754.  
CC HSP; Q01543; IFLI.  
CC InterPro; IPR000418; Ets.  
CC InterPro; IPR02341; HSF\_ETs.  
CC InterPro; IPR03118; SAM\_PNT.  
CC Pfam; PF00178; Ets; 1.  
CC Pfam; PF02198; SAM\_PNT; 1.  
CC PRINTS; PR00454; ETSDOMAIN.  
CC SMART; SM00413; ETS; 1.  
CC SMART; SM00251; SAM\_PNT; 1.  
CC PROSITE; PS00345; ETS\_DOMAIN\_1; 1.  
CC PROSITE; PS00346; ETS\_DOMAIN\_2; 1.  
CC PROSITE; PS50061; ETS\_DOMAIN\_3; 1.  
CC Transcription regulation; Activator; Nuclear protein; DNA-binding.  
KW DOMAIN 115 199 POINTED.  
FT DNA\_BIND 310 390 ETS-DOMAIN.  
FT SEQUENCE 478 AA; 53913 MW; 8DFB243ED8623A7 CRC64;

Query Match 49.4%; Score 42; DB 1; Length 478;  
Best Local Similarity 46.7%; Pred. No. 13;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 WVSQPPERTLEGSC 15  
DB 56 WLSQPPARVIRMEC 70

## RESULT 7

NXSL\_MICNI ID\_NXSL\_MICNI STANDARD; PRT; 60 AA.  
AC P80548;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Short neurotoxin 1 (Neurotoxin alpha).  
OS Micrurus nigrocinctus (Central American coral snake) (Gargantilla).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;  
OC Elapidae; Elapinae; Micrurus.  
OX NCBI\_TaxID=8635;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Venom;  
RX MEDLINE=96248443; PubMed=8665942;  
RA Rosso J.-P., Vargas-Rosso O., Gutierrez J.-M., Rochat H., Bougis P.E.;  
RT "Characterization of alpha-neurotoxin and phospholipase A2 activities  
RT from Micrurus venoms. Determination of the amino acid sequence and  
RT receptor-binding ability of the major alpha-neurotoxin from Micrurus  
RT nigrocinctus nigrocinctus.";  
RL Eur. J. Biochem. 238:231-239(1996).  
CC -1- FUNCTION: Produces peripheral paralysis by blocking neuromuscular  
CC transmission at the postsynaptic site. Binds to the nicotinic  
CC acetylcholine receptor.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Expressed by the venom gland.  
CC -1- SIMILARITY: Belongs to the snake toxin family.  
CC PIR; S68769; S68769.  
CC HSP; P01426; INEA.  
CC InterPro; IPR003571; Snake\_toxin.  
CC Pfam; PF00087; toxin; 1.  
CC ProDom; PD000206; Snake\_toxin; 1.  
CC PROSITE; PS00272; SNAKE\_TOXIN; 1.  
KW Toxin; Neurotoxin; Postsynaptic neurotoxin.  
FT DISULFID 3 22  
FT DISULFID 17 39 BY SIMILARITY.  
FT DISULFID 41 52 BY SIMILARITY.  
FT DISULFID 53 58 BY SIMILARITY.  
SQ SEQUENCE 60 AA; 5582 MW; 4E9580F886F0279C CRC64;

Query Match 48.8%; Score 41.5; DB 1; Length 60;  
Best Local Similarity 64.3%; Pred. No. 1.8;  
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 3 SQPPERTL-EGSC 15  
DB 9 SQPPTIKTCEGQC 22

## RESULT 8

YTXD\_BACSU ID\_YTXD\_BACSU STANDARD; PRT; 272 AA.  
AC P39063;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein ytxd.  
GN YTXD OR BSU29730.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=95020526; PubMed=7934817;  
RA Grundy F.J., Waters D.A., Takova T.Y., Henkin T.M.;  
RT "Identification of genes involved in utilization of acetate and  
RT acetoin in Bacillus subtilis.";  
RL Mol. Microbiol. 10:259-271(1993).

[2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98048467; PubMed=9387221;  
 RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;  
 RT "Sequencing and functional annotation of the *Bacillus subtilis* genes  
 in the 200 kb rnb-dnaB region.";  
 RL Microbiology 143:3431-3441(1997).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessières P., Bolotin A., Borchert S.,  
 RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medvedev V.,  
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rev M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zunstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.";  
 RL Nature 390:249-256(1997).  
 CC -1- FUNCTION: May be involved in some transport function.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the *motA* family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: LJ7309; AAC68282.1; --  
 CC EMBL: AF008220; AAC00300.1; --  
 CC EMBL: Z99119; CAB14951.1; --  
 CC PIR: S39641; S39641.  
 CC Subtilist; BG10365; ytxD.  
 CC InterPro: IPR000540; Flag MotA.  
 CC InterPro: IPR002898; MotA\_Exb.  
 CC Pfam: PF01618; MotA\_Exb; 1.  
 CC PROSITE: PS01307; MOTA; 1.  
 CC TRANSMEM 9 29 POTENTIAL.  
 CC TRANSMEM 38 58 POTENTIAL.  
 CC TRANSMEM 154 174 POTENTIAL.  
 CC TRANSMEM 188 208 POTENTIAL.  
 CC TRANSMEM 209 272 POTENTIAL.  
 CC DOMAIN 272 AA; 30143 MW; 376012B4BDA7490 CRC64;  
 CC SEQUENCE 272 AA; 30143 MW; 376012B4BDA7490 CRC64;  
 CC  
 CC Query Match 48.2%; Score 41; DB 1; Length 272;  
 CC Best Local Similarity 46.7%; Pred. No. 14;  
 CC Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Best Local Similarity 42.9%; Pred. No. 11;  
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 Qy 1 WVSQPEIRTEGSC 14  
 Db 255 WQKQPKQVTKKGS 268  
 RESULT 9  
 FPPS NEUCR STANDARD; PRT; 347 AA.  
 AC Q92250;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Farnesyl pyrophosphate synthetase (fpp synthetase) (FPP) (Farnesyl  
 DE pyrophosphate synthetase) [Includes: Dimethylallyltransferase  
 DE (EC 2.5.1.1); Geranyltransferase (EC 2.5.1.10)].  
 GN FPPS.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 CX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=74-OR23-1A / FGSC 987;  
 RX MEDLINE=96337905; PubMed=8753652;  
 RA Homann V., Mendez K., Arntz C., Ilardi V., Macino G., Morelli G.,  
 RA Bose G., Tudzynski B.,  
 RT "The isoprenoid pathway: cloning and characterization of fungal FPPS  
 genes.";  
 RL Curr. Genet. 30:232-239(1996).  
 CC -1- FUNCTION: Catalyzes the sequential condensation of isopentenyl  
 CC pyrophosphate with the allylic pyrophosphates, dimethylallyl  
 CC pyrophosphate, and then with the resultant geranylpyrophosphate  
 CC to the ultimate product farnesyl pyrophosphate.  
 CC -1- CATALYTIC ACTIVITY: Dimethylallyl diphosphate + isopentenyl  
 CC diphosphate = diphosphate + geranyl diphosphate.  
 CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate + isopentenyl diphosphate  
 CC = diphosphate + trans,trans-farnesyl diphosphate.  
 CC -1- PATHWAY: Isoprene biosynthesis, cholesterol biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to the FPP/GGPP synthetase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: X96944; CAA65645.1; --  
 CC PIR: S71436; S71436.  
 CC HSP: P08836; IUBX.  
 CC InterPro: IPR000092; Polyrenyl synt.  
 CC InterPro: IPR008949; Terpenoid synth.  
 CC Pfam: PF00348; polyrenyl synt; 1.  
 CC PROSITE: PS00444; POLYPRENYL SYNTHET 2; 1.  
 CC PROSITE: PS00723; POLYPRENYL SYNTHET 1; 1.  
 CC TRANSFERASE; Isoprene biosynthesis; Cholesterol biosynthesis.  
 CC ACT SITE 184 184 BY SIMILARITY.  
 CC SEQUENCE 347 AA; 40322 MW; 0D353712FB889ECE CRC64;  
 CC  
 CC Query Match 48.2%; Score 41; DB 1; Length 347;  
 CC Best Local Similarity 46.7%; Pred. No. 14;  
 CC Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 Qy 1 WVSQPEIRTEGSC 15  
 Db 37 WYKQSLVNTLGKGC 51

RESULT 10  
 LFTB MOUSE  
 ID LFTB\_MOUSE STANDARD; PRT; 368 AA.  
 AC P57785;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Left-right determination factor B precursor (Lefty-2 protein).  
 GN LFTB OR LEFTY2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98006264; PubMed=9348041;  
 RA Meno C., Ito Y., Saijoh Y., Matsuda Y., Tashiro K., Kuhara S.,  
 RA Hamada H.;  
 RT "Two closely-related left-right asymmetrically expressed genes, lefty-  
 RT 1 and lefty-2: their distinct expression domains, chromosomal linkage  
 RT and direct neutralizing activity in Xenopus embryos.";  
 RL Genes Cells 2:513-524 (1997).  
 CC -1- FUNCTION: MORPHOGEN FOR LEFT-RIGHT ASYMMETRY DETERMINATION OF  
 CC ORGAN SYSTEMS IN MAMMALS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC EXPRESSED IN THE EMERGING MESODERM. BY ES.0, EXPRESSED EXCLUSIVELY  
 CC ON THE LEFT SIDE OF DEVELOPING MESODERM. BY ES.0, EXPRESSED EXCLUSIVELY  
 CC PREDOMINANTLY IN THE LATERAL-PLATE MESODERM (LPM). WEAK EXPRESSION  
 CC IN THE PROSPECTIVE FLOOR PLATE (FPF).  
 CC -1- PTM: The processing of the protein may also occur at the second R-  
 CC X-X-R site located at AA 132-135. Processing appears to be  
 CC regulated in a cell-type specific manner.  
 CC -1- SIMILARITY: Belongs to the TGF-beta family.  
 DR HSP; P10600; ITGJ.  
 DR MGD; MGI:107405; Leftb.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 DR Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal;  
 KW Multigene family.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT PROPEP 22 77 OR 135 (POTENTIAL).  
 FT CHAIN 78 368 LEFT-RIGHT DETERMINATION FACTOR B.  
 FT DISULFID 253 266 BY SIMILARITY.  
 FT DISULFID 265 318 BY SIMILARITY.  
 FT DISULFID 295 353 BY SIMILARITY.  
 FT DISULFID 299 355 BY SIMILARITY.  
 FT CARBOHYD 158 158 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 368 AA; 41141 MW; 56B5ED095167A6C3 CRC64;  
 Query Match 48.2%; Score 41; DB 1; Length 368;  
 Best Local Similarity 47.1%; Pred. No. 15;  
 Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;  
 QY 1 WVSQPEIRTE--GSC 15  
 | : : : : :  
 Db 283 WILEPPGLTYECVGC 299  
 RESULT 11  
 TGF4 MOUSE  
 ID TGF4\_MOUSE STANDARD; PRT; 368 AA.  
 AC Q64280;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transforming growth factor beta 4 precursor (TGF-beta 4) (Lefty  
 DE protein) (Lefty-1 protein) (STRA3 protein).

GN EBAF OR TGFb4 OR STRA3 OR LEFTY OR LEFTY1.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96202359; PubMed=8610011;  
 RA Meno C., Saijoh Y., Fujii H., Ikeda M., Yokoyama T., Yokoyama M.,  
 RA Toyoda Y., Hamada H.;  
 RT "Left-right asymmetric expression of the TGF beta-family member lefty  
 RT in mouse embryos.";  
 RL Nature 381:151-155 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX Bouillet P.;  
 RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98156497; PubMed=9496783;  
 RA Oulad-Abdelghani M., Chazaud C., Bouillet P., Mattei M.-G., Dolle P.,  
 RA Chambon P.;  
 RT "Stra3/lefty, a retinoic acid-inducible novel member of the  
 RT transforming growth factor-beta superfamily.";  
 RL Int. J. Dev. Biol. 42:23-32 (1998).  
 RN [4]  
 RP FUNCTION.  
 MEDLINE=98372436; PubMed=9708731;  
 RA Meno C., Shimon A., Saijoh Y., Yashiro K., Mochida K., Ohishi S.,  
 RA Noji S., Kondoh H., Hamada H.;  
 RT "Lefty-1 is required for left-right determination as a regulator of  
 RT lefty-2 and nodal";  
 RL Cell 94:287-297 (1998).  
 CC -1- FUNCTION: REQUIRED FOR LEFT-RIGHT AXIS DETERMINATION AS A  
 CC REGULATOR OF LEFT2 AND NODAL.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DEVELOPMENTAL STAGE: BY ES.0, EXPRESSED EXCLUSIVELY ON THE LEFT  
 CC SIDE OF DEVELOPING EMBRYOS WITH EXPRESSION PREDOMINANTLY IN THE  
 CC PROSPECTIVE FLOOR PLATE (FPF). WEAK EXPRESSION IN THE LATERAL-  
 CC PLATE MESODERM (LPM).  
 CC -1- PTM: The processing of the protein may also occur at the second R-  
 CC X-X-R site located at AA 132-135. Processing appears to be  
 CC regulated in a cell-type specific manner.  
 CC -1- SIMILARITY: Belongs to the TGF-beta family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; D83921; BAA1212.1; -;  
 DR EMBL; Z73151; CAA97497.1; -;  
 DR EMBL; AJ000082; CAA03909.1; -;  
 DR EMBL; AJ000083; CAA03910.1; -;  
 DR PIR; S67507; S67507.  
 DR HSP; P10600; ITGJ.  
 DR MGD; MGI:107405; EbaF.  
 DR InterPro; IPR001839; TGFb.  
 DR InterPro; IPR001111; TGFb\_N.  
 DR Pfam; PF00019; TGF-beta; 1.  
 DR Pfam; PF00688; TGF-beta; 1.  
 DR ProDom; PD000357; TGFb; 1.  
 DR SMART; SM00204; TGFb; 1.  
 DR PROSITE; PS00250; TGF\_BETA\_1; 1.  
 DR Developmental protein; Growth factor; Cytokine; Glycoprotein; Signal.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 76 OR 135 (POTENTIAL).  
 FT CHAIN 77 368 TRANSFORMING GROWTH FACTOR BETA 4.  
 FT DISULFID 253 266 BY SIMILARITY.  
 FT DISULFID 265 318 BY SIMILARITY.



```
FT DISULFID 295 353 BY SIMILARITY.
FT DISULFID 299 355 BY SIMILARITY.
FT CARBOHYD 158 158 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 368 AA; 41497 MW; 821DAE663C546B5F CRC64;

Query Match 48.2%; Score 41; DB 1; Length 368;
Best Local Similarity 47.1%; Pred. No. 15;
Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 1 WVSQPEIRLE-GSC 15
   : : : : :
Db 283 WLEPPGFLTYECVGC 299

RESULT 12
EOMD_HUMAN STANDARD; PRT; 686 AA.
AC 095936; Q8TAZ2; Q9UPM7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eomesodermin homolog.
GN EOMES OR TB2
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99337662; PubMed=10407135;
RA Kimura N., Nakashima K., Ueno M., Taga T.;
RT "A novel mammalian T-box-containing gene, Tbx2, expressed in mouse
RT developing brain."
RL Brain Res. Dev. Brain Res. 115:183-193(1999).
RN [2]
RP SEQUENCE OF 291-455 FROM N.A.
RX MEDLINE=99107806; PubMed=9888994;
RA Yi C.-H., Terrett J.A., Li Q.-F., Ellington K., Packham E.A.,
RA Armstrong-Buisseret L., McClure P., Slingsby T., Brook J.D.;
RT "Identification, mapping and phylogenomic analysis of four new human
RT members of the T-box gene family: EOMES, TBX6, TBX18, and TBX19."
RL Genomics 55:10-20(1999).
RN [3]
RP SEQUENCE OF 297-686 FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marushina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellon E., Kettman M., Kuehnen M., Shevchenko Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SURCELLULAR LOCATION: Nuclear (Potential).
CC -!- SIMILARITY: Contains 1 T-box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB031038; BA383417.1; -
DR EMBL; AJ010280; CA37939.1; -
DR EMBL; BC025363; AA25363.1; -
DR HSSP; P24781; 1XBR.
DR TRANSFAC; T04395; -
DR Genew; HGNC:3372; EOMES.
DR MIM; 604615; -
DR GO; GO:0003677; F:DNA binding; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR InterPro; IPR008967; P53-like.
DR InterPro; IPR001699; TF_T-box.
DR Pfam; PF00907; T-box; 1.
DR PRINTS; PR00937; TBOX.
DR SMART; SMO0425; TBOX; 1.
DR PROSITE; PS01283; TBOX_1; 1.
DR PROSITE; PS01264; TBOX_2; 1.
DR PROSITE; PS0252; TBOX_3; 1.
KW Developmental protein; Transcription regulation; DNA-binding;
KW Nuclear protein.
KW DOMAIN 117 132 POLY-ALA.
FT DOMAIN 223 229 POLY-GLY.
FT DOMAIN 246 249 POLY-ALA.
FT DNA BIND 276 456 T-BOX.
FT DOMAIN 381 384 POLY-ASN.
FT CONFLICT 337 337 C -> S (IN REF. 2).
FT CONFLICT 411 411 E -> EK (IN REF. 2).
FT CONFLICT 415 415 E -> D (IN REF. 2).
FT CONFLICT 661 661 G -> E (IN REF. 3).
FT CONFLICT 685 685 S -> T (IN REF. 3).
SQ SEQUENCE 686 AA; 72612 MW; 91A9402B1C6265FD CRC64;

Query Match 48.2%; Score 41; DB 1; Length 686;
Best Local Similarity 35.7%; Pred. No. 29;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 WVSQPEIRLEGS 14
   : : : : :
Db 618 WIEPPSIKSLDSN 631

RESULT 13
Y297_HUMAN STANDARD; PRT; 1411 AA.
AC Q15040; Q9UEG6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein KIAA0297/KIAA0329.
GN KIAA0297 OR KIAA0329.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT the complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 4:141-150(1997).
CC -!- SIMILARITY: TO HUMAN KIAA1017.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AB002327; BAA20787.1; -  
 DR EMBL: AB002295; BAA20757.1; -  
 DR InterPro: IPR006624; TECPR.  
 DR InterPro: IPR001680; WD40.  
 DR SMART; SMO0706; TECPR; 10.  
 DR SMART; SMO0320; WD40; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1411 AA; 153861 MW; B694CA9CB6124AD2 CRC64;

Query Match 48.2%; Score 41; DB 1; Length 1411;  
 Best Local Similarity 50.0%; Pred. No. 62;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WVSQPPPEIRTELEGS 14  
 | : ||| : |||  
 Db 739 WLFQPRDQLTSS 752

RESULT 14  
 WBBJ\_ECOLI STANDARD; PRT; 196 AA.  
 ID WBBJ\_ECOLI  
 AC P37750; P76375;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative lipopolysaccharide biosynthesis O-acetyl transferase wbbJ  
 DE (EC 2.3.1.-).  
 GN WBBJ OR B2033.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=94292434; PubMed=7517390;  
 RA Yao Z., Valvano M.A.;  
 RT "Genetic analysis of the O-specific lipopolysaccharide biosynthesis  
 RT region (rfb) of Escherichia coli K-12 W3110: identification of genes  
 RT that confer group 6 specificity to Shigella flexneri serotypes Y and  
 RT 4a";  
 RL J. Bacteriol. 176:4133-4143(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / WGI;  
 RX MEDLINE=94292435; PubMed=7517391;  
 RA Stevenson G., Neal B., Liu D., Hobbs M., Packer N.H., Batley M.,  
 RA Redmond J.W., Lindquist L., Reeves P.R.;  
 RT "Structure of the O antigen of Escherichia coli K-12 and the sequence  
 RT of its rfb gene cluster";  
 RN J. Bacteriol. 176:4144-4156(1994).  
 RN [3]  
 RP REVISIONS TO 168-176 AND 187-189.  
 RC STRAIN=K12 / WGI;  
 RA Stevenson G.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / WGI655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1232-1244(1997).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;

REX MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,  
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,  
 RA Saito N., Sampei G., Seki Y., Sivasubramanian S., Tagami H.,  
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;  
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 40.1-50.0 min region on the linkage map";  
 RL DNA Res. 3:379-392(1996).  
 CC -!- FUNCTION: PUTATIVE O-ACETYLTRANSFERASE THAT TRANSFERS AN O-ACETYL  
 CC ON THE O ANTIGEN.  
 CC -!- PATHWAY: Lipopolysaccharide biosynthesis.  
 CC -!- SIMILARITY: BELONGS TO THE CYSE/LACA/LPXA/NOFL FAMILY OF  
 CC ACETYLTRANSFERASES. COMPOSED OF MULTIPLE REPEATS OF [LIV]-G-X(4).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U03041; AAC31635.1; -  
 DR EMBL: U03876; AAB88406.1; -  
 DR EMBL: AE000294; AAC75094.1; -  
 DR EMBL: D90841; BAA15875.1; -  
 DR EMBL: D90842; BAA15887.1; -  
 DR PIR; H64968; H64968.  
 DR EcoGene; EG11984; wbbJ.  
 DR InterPro: IPR001451; Hexapep\_transf.  
 DR Pfam; PF00132; hexapep; 4.  
 DR PROSITE; PS00101; HEXAPEP TRANSFERASES; FALSE NEG.  
 KW Lipopolysaccharide biosynthesis; Transferase; Acyltransferase; Repeat;  
 KW Complete proteome.  
 FT CONFLICT 168 176 SIPENTVIA -> LPRKYCHC (IN REF. 1).  
 FT CONFLICT 187 189 NHE -> IMR (IN REF. 1).  
 SQ SEQUENCE 196 AA; 21675 MW; DIC2FAD7B29A1B1 CRC64;

Query Match 47.1%; Score 40; DB 1; Length 196;  
 Best Local Similarity 70.0%; Pred. No. 12;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 PPPIRTLEGS 14  
 | : ||| : |||  
 Db 122 PPDRTLESS 131

RESULT 15  
 CREB\_HUMAN STANDARD; PRT; 341 AA.  
 ID CREB\_HUMAN  
 AC P16220; P21934; Q9DMA7;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE cAMP response element binding protein (CREB).  
 GN CREB1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM CREB-A).  
 RX MEDLINE=90319091; PubMed=2142528;  
 RA Berkowitz L.A., Gilman M.Z.;  
 RT "Two distinct forms of active transcription factor CREB (cAMP  
 RT response element binding protein)";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5258-5262(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM CREB-A).  
 RX MEDLINE=90316112; PubMed=2196176;  
 RA Yoshimura T., Fujisawa J.I., Yoshida M.;

RT Multiple cDNA clones encoding nuclear proteins that bind to the tax-  
RT dependent enhancer of HTLV-1; all contain a leucine zipper structure  
RT and basic amino acid domain.";  
RL ENBO J. 9:2537-2542(1990).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM CREB-A).  
RX MEDLINE=92087371; PubMed=1966745;  
RA Weeber G., Meyer T.E., Hoeffler J.P., Habener J.F.;  
RT "Diversification of cyclic AMP-responsive enhancer binding proteins-  
RT generated by alternative exon splicing.";  
RL Trans. Assoc. Am. Physicians 103:28-37(1990).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM CREB-B).  
RX MEDLINE=9072714; PubMed=2974179;  
RA Hoeffler J.P., Meyer T.E., Yun Y., Jameson J.L., Habener J.F.;  
RT "Cyclic AMP-responsive DNA-binding protein: structure based on a  
RT cloned placental cDNA.";  
RL Science 242:1430-1433(1988).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM CREB-B).  
RX MEDLINE=91334144; PubMed=1831258;  
RA Short M.L., Manohar C.F., Furtado M.R., Ghadge G.D., Wolinsky S.M.,  
RT Thimmapaya B., Jungmann R.A.;  
RL "Nucleotide and derived amino-acid sequences of the CRE-binding  
RT proteins from rat C6 glioma and HeLa cells.";  
RL Nucleic Acids Res. 19:4290-4290(1991).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM CREB-A).  
RX TISSUE=Eye;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RT Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Ziegler B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Spletton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RL human and mouse cDNA sequences.";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7]  
RP SEQUENCE OF 1-8 FROM N.A.  
RX MEDLINE=93145890; PubMed=8381074;  
RA Meyer T.E., Weeber G., Lin J., Beckmann W., Habener J.F.;  
RT "The promoter of the gene encoding 3' 5'-cyclic adenosine  
RT monophosphate (cAMP) response element binding protein contains  
RT cAMP response elements: evidence for positive autoregulation of gene  
RT transcription.";  
RL Endocrinology 132:770-780(1993).  
RN [8]  
RP FUNCTION: THIS PROTEIN BINDS THE CAMP RESPONSE ELEMENT (CRE), A  
CC SEQUENCE PRESENT IN MANY VIRAL AND CELLULAR PROMOTERS. CREB  
CC STIMULATES TRANSCRIPTION ON BINDING TO THE CRE.  
CC -1- SUBUNIT: Binds DNA as a dimer.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=CREB-A;  
CC IsoId=P16220-1; Sequence=Displayed;  
CC Name=CREB-B;  
CC IsoId=P16220-2; Sequence=VSP\_000596;  
CC -1- PM: Stimulated by phosphorylation.  
CC -1- SIMILARITY: Belongs to the bZIP family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; S72459; AAB20597.1; -  
DR EMBL; X55545; CAA39151.1; -  
DR EMBL; M34356; ABA35717.1; -  
DR EMBL; M34356; ABA35716.1; -  
DR EMBL; M27691; AAA35715.1; -  
DR EMBL; X60003; CAA42620.1; -  
DR EMBL; BC10636; AAH10636.1; -  
DR EMBL; S53724; AAD13869.1; -  
DR PIR; A37340; A35769;  
DR PIR; B37340; B35769;  
DR PIR; S22298; S22298;  
DR TRANSFAC; T00163; -  
DR TRANSFAC; T00166; -  
DR GENBANK; HGNC:2345; CREB1.  
DR MIM; 123810; -  
DR GO; GO:0005634; C:nucleus; TAS.  
DR GO; GO:0003712; F:transcription cofactor activity; TAS.  
DR GO; GO:0003700; F:transcription factor activity; TAS.  
DR GO; GO:0007165; P:signal transduction; TAS.  
DR GO; GO:0006351; P:transcription, DNA-dependent; TAS.  
DR InterPro; IPR001630; Leuzip\_CREB.  
DR InterPro; IPR003102; PKID.  
DR Pfam; PF00170; bZIP; 1.  
DR Pfam; PF02173; PKID; 1.  
DR PRINTS; PR00041; LEUZIPPRCREB.  
DR SMART; SM00338; BRLZ; 1.  
DR PROSITE; PS00217; bZIP; 1.  
DR PROSITE; PS00036; bZIP\_BASIC; 1.  
KW Transcription regulation; DNA-binding; Activator; Phosphorylation;  
KW Nuclear protein; Alternative splicing.  
FT DNA\_BIND 284 305 BASIC MOTIF.  
FT DOMAIN 311 332 LEUCINE-ZIPPER.  
FT VARSPLIC 88 101 Missing (in isoform CREB-B).  
FT CONFLICT 4 4 E -> D (IN REF. 5).  
FT CONFLICT 8 8 E -> D (IN REF. 5).  
FT CONFLICT 160 160 T -> A (IN REF. 5).  
FT CONFLICT 167 167 T -> A (IN REF. 5).  
FT CONFLICT 169 169 T -> A (IN REF. 5).  
FT CONFLICT 176 176 Q -> R (IN REF. 5).  
FT CONFLICT 184 184 A -> T (IN REF. 5).  
FT CONFLICT 188 188 G -> R (IN REF. 5).  
FT CONFLICT 195 195 N -> S (IN REF. 5).  
FT CONFLICT 210 210 T -> A (IN REF. 5).  
SQ SEQUENCE 341 AA; 36688 MW; D5E989AE40BF69AF CRC64;  
Query Match 47.1%; Score 40; DB 1; Length 341;  
Best Local Similarity 42.9%; Pred.No. 21;  
Matches 6; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
QY 2 VSQPEIRTLGGSC 15  
DB 77 VIQSPQVQTVOSSC 90  
Search completed: February 26, 2004, 12:10:54  
Job time : 3.27321 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:04:48 ; Search time 6.40584 Seconds  
(without alignments)  
738.822 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPPEIRTLGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	89.4	115	6 Q8MJ00	Q8mj00 macaca mula
2	76	89.4	151	6 Q8MJ01	Q8mj01 macaca mula
3	76	89.4	152	4 Q95668	Q95668 homo sapien
4	76	89.4	165	4 Q95669	Q95669 homo sapien
5	76	89.4	176	4 Q95667	Q95667 homo sapien
6	76	89.4	176	6 Q95JB8	Q95JB8 macaca fasc
7	76	89.4	177	4 Q95J30	Q95J30 macaca mula
8	76	89.4	180	6 Q8MJ02	Q8mj02 macaca mula
9	76	89.4	190	4 Q95J32	Q95J32 homo sapien
10	76	89.4	201	4 Q95J31	Q95J31 homo sapien
11	62	72.9	192	11 Q8CG11	Q8cg11 rattus norv
12	62	72.9	192	11 Q8QW88	Q8QW88 rattus norv
13	62	72.9	192	11 Q8CFD9	Q8cfD9 rattus norv
14	49	57.6	2201	5 Q8EN11	Q8en11 drosophila
15	49	57.6	2529	5 Q9VWC0	Q9VWC0 drosophila
16	48	56.5	1512	2 Q93HF2	Q93hf2 streptomyce

17	48	56.5	1517	16	Q826K7	Q826k7 streptomyce
18	46.5	54.7	155	16	Q92GC3	Q92gc3 rickettsia
19	46.5	54.7	155	16	Q92CH2	Q92ch2 rickettsia
20	46	54.1	236	8	Q9TCC7	Q9Tcc7 nephroselimi
21	46	54.1	397	10	Q84QL3	Q84ql3 oryza sativ
22	45	52.9	196	10	Q8H8H6	Q8h8h6 oryza sativ
23	45	52.9	353	10	Q80449	Q80449 arabidopsis
24	45	52.9	353	10	Q8LCL0	Q8lcl0 arabidopsis
25	44	51.8	650	16	Q8FWS2	Q8fws2 bruceella su
26	44	51.8	688	11	Q9JUL1	Q9jJl1 mus musculu
27	44	51.8	707	11	Q8EN22	Q8en22 mus musculu
28	43.5	51.2	804	4	Q8TER5	Q8ter5 homo sapien
29	43	50.6	305	16	Q9WX21	Q9wx21 streptomyce
30	43	50.6	456	13	Q9W6Z9	Q9w6z9 xenopus lae
31	43	50.6	474	2	Q9ADU2	Q9adu2 salmonella
32	43	50.6	485	13	Q9W700	Q9w700 xenopus lae
33	43	50.6	514	2	Q41111	Q41111 anabaena sp
34	43	50.6	514	16	Q8YWY3	Q8ywy3 anabaena sp
35	43	50.6	687	16	Q8YF69	Q8yf69 anabaena sp
36	42	49.4	101	16	Q9RX23	Q9rx23 deinococcus
37	42	49.4	360	11	Q9D6W8	Q9d6w8 mus musculu
38	42	49.4	451	13	Q8UUU0	Q8uuu0 gallus gall
39	42	49.4	455	11	Q91XV5	Q91xv5 rattus norv
40	42	49.4	479	4	Q8IXK9	Q8ixk9 homo sapien
41	42	49.4	535	16	P73052	P73052 synchocyst
42	42	49.4	644	4	Q9NRK6	Q9nrk6 homo sapien
43	42	49.4	976	16	Q8FTA0	Q8fta0 corynebacte
44	42	49.4	1317	16	Q910F4	Q910f4 pseudomonas
45	41	48.2	91	10	Q84VG8	Q84vg8 arabidopsis

#### ALIGNMENTS

RESULT 1  
Q8MJ00 PRELIMINARY; PRT; 115 AA.  
AC Q8MJ00; 01-OCT-2002 (T-EMBLrel. 22, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (T-EMBLrel. 25, Last annotation update)  
DE Nkp30S.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
OC Cercopitheciinae; Macaca.  
NCBI\_TaxID=9544;  
[1]  
RP SEQUENCE FROM N.A.  
RA LaBonte M.L.; Miller J.; Letvin N.L.;  
RT "Molecular cloning of rhesus monkey Nkp46 and Nkp30 and identification of Nkp46SD and Nkp30S."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035217; AXK63119.1; -  
DR InterPro; IPR003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR Pfam; PF00047; IG; 1.  
DR SMART; SM00409; IG; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
SQ SEQUENCE 115 AA; 12871 MW; 1D34CC0E8986DE9F CRC64;  
Query Match 89.4%; Score 76; DB 6; Length 115;  
Best Local Similarity 100.0%; Pred. No. 1.5e-05;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WVSQPPPEIRTLGGS 14  
|||  
DB 20 WVSQPPPEIRTLGGS 33  
|||  
RESULT 2  
Q8MJ01 PRELIMINARY; PRT; 151 AA.  
ID Q8MJ01

AC QBMJ01;  
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE NKp30v1ED  
 OS Macaca mulatta (Rhesus macaque).  
 OC Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Cercopitheciinae; Macaca.  
 OC NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA LaBonte M.L.; Miller J.; Letvin N.L.;  
 RT "Molecular cloning of rhesus monkey NKp46 and NKp30 and identification  
 RT of NKp46SD and NKp30S.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY035216; AAK63118.1; --  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; ig; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS00835; IG\_LIKE; 1.  
 SQ SEQUENCE 151 AA; 16372 MW; 41E1FD771DB70918 CRC64;  
 Query Match 89.4%; Score 76; DB 6; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 2e-05; 0; Indels 0; Gaps 0;  
 Matches 14; Conservative 0; Mismatches 0;  
 QY 1 WVSQPPPIRTLEGS 14  
 DB |||||  
 20 WVSQPPPIRTLEGS 33  
 RESULT 3  
 ID 095668 PRELIMINARY; PRT; 152 AA.  
 AC 095668;  
 DT 01-MAY-1999 (T-EMBLrel. 10, Created)  
 DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE 1C7d.  
 GN 1C7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93272029; PubMed=8499947;  
 RA Iris F.; Bougueleret L.; Prieur S.; Caterina D.; Primas G.; Perrot V.;  
 RA Jurka J.; Rodriguez-Tome P.; Clavierie J.; Cohen D.; Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NFKAPPAB  
 RT family within a 90 kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96215741; PubMed=8629302;  
 RA Utans U.; Quist W.C.; McManus B.M.; Wilson J.E.; Arcoci R.J.;  
 RA Wallace A.F.; Russell M.E.;  
 RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I.; de Baey A.; Messer G.; Kick G.; Zwierrina H.;  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=93208881; PubMed=7916655;

RA Browning J.L.; Ngam-ek A.; Lawton P.; DeMarinis J.; Tizard R.;  
 RA Chow E.P.; Hession C.; O'Brine-Greco B.; Foley S.F.; Ware C.F.;  
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E.; Naylor S.L.; Sakaguchi A.Y.; Smith D.; Jarrett-Nedwin J.;  
 RA Pennica D.; Goeddel D.V.; Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91086846; PubMed=1670638;  
 RA Messer G.; Spengler U.; Jung M.C.; Honold G.; Bloemer K.; Pape G.R.;  
 RA Rethmuelier G.; Weiss E.H.;  
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
 RT Level of TNF-beta Production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91139175; PubMed=1671667;  
 RA Abraham I.J.; Du D.C.; Zahedi K.; Dawkins R.L.; Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53(1991).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94382679; PubMed=8081366;  
 RA Albertella M.R.; Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L.; Chardon P.; Nunes M.; Renard C.; Geffrotin C.; Vaiman M.;  
 RA Van Zeveren A.; Coppieters W.; Van de Weghe A.; Bouquet Y.; Choy W.;  
 RA Strominger J.; Spies T.;  
 RT "The B2M1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
 RT Nuclear RNA Helicase of the D-E-A-D Family.";  
 RL Genomics 26:210-218(1995).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J.; Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A.; Fellerhoff B.; Maier S.; Martinuzzi S.; Weidle U.;  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing.";  
 RL Genomics 45:591-600(1997).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=98149985; PubMed=9480751;  
 RA Shiina T.; Tamiya G.; Oka A.; Yamagata T.; Yamagata N.; Kikkawa E.;  
 RA Goto K.; Mizuki N.; Watanabe K.; Fukuzumi Y.; Taguchi S.; Sugawara C.;  
 RA Ono A.; Chen L.; Yamazaki M.; Tashiro H.; Ando A.; Ikemura T.;  
 RA Kimura M.; Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT LKB1 and MICA genes at the centromeric end of the HLA class I  
 RL region.";  
 RL Genomics 47:372-382(1998).  
 DR EMBL; Y14768; CAA75067.1; --

DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00403; IG; 1.  
 DR PROSITE; PS00835; IG-LIKE; 1.  
 SQ SEQUENCE 152 AA; 16393 MW; 42718746451F9ADC CRC64;

Query Match 89.4%; Score 76; DB 4; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 2e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPERTLEGS 14  
 |||||  
 DB 20 WVSQPPERTLEGS 33  
 |||||

RESULT 4

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bouguet-Lemerle L., Prieur S., Caterina D., Primas G., Perrot V.,  
 Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.,  
 "Dense Alu clustering and a potential new member of the NFKappaB  
 family within a 90 Kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bouguet-Lemerle L., Prieur S., Caterina D., Primas G., Perrot V.,  
 Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.,  
 "Dense Alu clustering and a potential new member of the NFKappaB  
 family within a 90 Kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcenci R.J.,  
 Wallace A.F., Russell M.E.;  
 RA "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 Weiss E.H.;  
 RA "Cloning and genomic characterization of LST1: a new gene in the human  
 TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RA "Lymphotoxin-beta: A new member of the TNF family that forms a  
 heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 Pennica D., Goeddel D.V., Gray P.W.;  
 RA "Human lymphotoxin and tumor necrosis factor genes: structure,  
 homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).

RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 Riethmuller G., Weiss E.H.;  
 RA "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
 Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
 Correlates with A Variant Amino Acid in Position 26 and a Reduced  
 Level of TNF-beta Production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RA "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53(1991).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RA "Characterization of a novel gene in the human major  
 histocompatibility complex that encodes a potential new member of the  
 I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
 Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 Strominger J., Spies T.;  
 RA "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
 Nuclear RNA Helicase of the D-E-A-D Family.";  
 RL Genomics 26:210-218(1995).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RA "Alternative splicing of the LST-1 gene located in the major  
 histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
 Weiss E.H.;  
 RA "Complex expression pattern of the TNF region gene LST1 through  
 differential regulation, initiation, and alternative splicing.";  
 RL Genomics 45:591-600(1997).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
 Kimura M., Inoko H.;  
 RA "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 LxBL and MICA genes at the centromeric end of the HLA class I  
 region.";  
 RL Genomics 47:372-382(1998).  
 DR EMBL; Y14768; CAZ5068.1;  
 DR GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR GO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00403; IG; 1.  
 DR PROSITE; PS00835; IG-LIKE; 1.  
 SQ SEQUENCE 165 AA; 17796 MW; 0A3CA691CA3E1E7E CRC64;

Query Match 89.4%; Score 76; DB 4; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1: WVSQPPEIRTELEGS 14  
 Db |||||

## RESULT 5

ID O95667 PRELIMINARY; PRT; 176 AA.  
 AC O95667;  
 DT 01-MAY-1999 (T-EMBLrel. 10, Created)  
 DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE 1C7e.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9372029; PubMed=8499947;  
 RA Iris F., Bouguieret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tome P., Claverie J., Cohen D., Dausset J.,  
 RT "Dense Alu clustering and a potential new member of the NFKAPPAB  
 family within a 90 kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
 molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Graco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin-beta: A new member of the TNF family that forms a  
 heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Rietmueller G., Weiss E.H.;  
 RT "Polymorphic Structure of the Tumor Necrosis Factor (TNF) Locus: An  
 RT Ncol Polymorphism in the First Intron of the Human TNF-beta Gene  
 RT Correlates with A Variant Amino Acid in Position 26 and a Reduced  
 RT Level of TNF-beta Production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";

Immunogenetics 33:50-53(1991).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vaiman M.,  
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The BAT1 Gene in the MHC Encodes an Evolutionarily Conserved Putative  
 RT Nuclear RNA Helicase of the D-E-A-D Family.";  
 RL Genomics 26:210-218(1995).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing.";  
 RL Genomics 45:591-600(1997).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98149985; PubMed=9480751;  
 RA Shilina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando A., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IKB and MICA genes at the centromeric end of the HLA class I  
 RT region.";  
 RL Genomics 47:372-382(1998).  
 DR EMBL: Y14768; CAA75066.1;  
 DR GO: GO:0003793; P:defense/immunity protein activity; NAS.  
 DR GO: GO:0006955; P:immune response; NAS.  
 DR GO: GO:0006954; P:inflammatory response; NAS.  
 DR InterPro: IPR003599; IG.  
 DR InterPro: IPR007110; IG-like.  
 DR Pfam: PF00047; IG; 1.  
 DR SMART: SM00409; IG; 1.  
 DR PROSITE: PS50835; IG LIKE; 1.  
 SQ SEQUENCE 176 AA; 18749 MW; 162BBB775DA2BCD35 CRC64;  
 Query Match 89.4%; Score 76; DB 4; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WVSQPPEIRTELEGS 14  
 Db |||||

## RESULT 6

Q95JB8 PRELIMINARY; PRT; 176 AA.  
 ID O95JB8  
 AC O95JB8  
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE NFKB30 (NFKB30v1).  
 GN NCR.



OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541, 9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.fascicularis; TISSUE=Lymphoid;  
 RA Rizzi M., Biassoni R.;  
 RT "Non MHC specific Natural cytotoxicity receptors (NCR) expressed in  
 RT Macaca fascicularis lymphoid cells.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.mulatta;  
 RA LaBonte M.L., Miller J., Letvin N.L.;  
 RT "Molecular cloning of rhesus monkey NKp46 and identification  
 RT of NKp46SD and NKp30S.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ273899; CAC41081.1; -;  
 DR EMBL; AY035215; AAK63117.1; -;  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 SQ SEQUENCE 176 AA; 19251 MW; 97B2A3B625E4AD54 CRC64;  
 Query Match 89.4%; Score 76; DB 6; Length 176;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WVSQPPEINTLEGS 14  
 Db 20 WVSQPPEINTLEGS 33  
 RESULT 7  
 OI4930  
 ID OI4930 PRELIMINARY; PRT; 177 AA.  
 AC OI4930;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE 1C7 precursor (1C7 protein).  
 GN 1C7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=96422187; PubMed=8924804;  
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;  
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";  
 RL Genomics 31:215-222(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX Nalabolu S.R., Raghunathan A., Weissman S.M.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Irls F., Bougueret L., Prieur S., Caterina D., Perrot V.,  
 RA Jurka J., Rodriguez-Tone P., Claverie J., Cohen D., Dausset J.;  
 RT "Dense Alu clustering and a potential new member of the NF kappa B  
 RT family within a 90 kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arcaci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392(1996).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierrina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region.";  
 RL Immunogenetics 42:315-322(1995).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53(1991).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Feelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vainan M.,  
 RA Van Zeveren A., Coppieters W., Van de Waghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family.";  
 RL Genomics 26:210-218(1995).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;  
 RA de Baey A., Feilerhoff B., Maier S., Martinozzi S., Weidie U.,  
 RA Weiss E.H.;



RT "Complex expression pattern of the TNF region gene LST1 through  
 RL differential regulation, initiation, and alternative splicing."  
 RL Genomics 45:591-600(1997).  
 RN [14]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=9814985; PubMed=9480751;  
 RA Shiina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Teguchi S., Sugawara C.,  
 RA Oka A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikenura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IRL and MICA genes at the centromeric end of the HLA class I  
 RT region."  
 RL Genomics 47:372-382(1998).  
 RN [15]  
 RN SEQUENCE FROM N.A.  
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [16]  
 RN SEQUENCE FROM N.A.  
 RA Shiina S., Tamiya G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF031136; AAB86578.1; -;  
 DR EMBL; Y14768; CAAV5064.1; -;  
 DR EMBL; AP00505; BAB63393.1; -;  
 DR Genew; HGNC:14189; LST1.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0003793; F:defense/immunity protein activity; TAS.  
 DR GO; GO:0006968; P:cellular defense response; TAS.  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW SIGNAL.  
 FT STGNAL.  
 FT CHAIN  
 FT SEQUENCE 177 AA; 19237 MW; DD5EC96F0AB2DCE6 CRC64;  
 SEQUENCE 177 AA; 19237 MW; DD5EC96F0AB2DCE6 CRC64;  
 Query Match 89.4%; Score 76; DB 4; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WVSQPPPIRTLEGS 14  
 Db 20 WVSQPPPIRTLEGS 33  
 RESULT 8  
 ID QBMJ02 PRELIMINARY; PRT; 180 AA.  
 AC QBMJ02;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Nkp30.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 ON NCBI\_TaxID=9544;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA LaBonte M.L., Miller J., Letvin N.L.;  
 RT "Molecular cloning of rhesus monkey Nkp46 and identification  
 RT of Nkp46SD and Nkp30S."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY035214; AAK63116.1; -;  
 DR InterPro; IPR003599; IG.  
 DR InterPro; IPR007110; IG-like.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SM00409; IG; 1.

DR PROSITE; PS50835; IG\_LIKE; 1.  
 SQ SEQUENCE 180 AA; 19639 MW; 5C2DF53487B2A3B6 CRC64;  
 Query Match 89.4%; Score 76; DB 6; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WVSQPPPIRTLEGS 14  
 Db 20 WVSQPPPIRTLEGS 33  
 RESULT 9  
 ID O14932 PRELIMINARY; PRT; 190 AA.  
 AC O14932;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE IC7 precursor (Natural killer cell receptor).  
 GN IC7 OR NKP30.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=96422187; PubMed=8824804;  
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;  
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC."  
 RL Genomics 31:215-222(1996).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RA Nalabolu S.R., Raghunathan A., Weissman S.M.;  
 RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bougueleret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodrigues-Tome P., Claverie J., Cohen D., Dausset J.;  
 RT "Jense Alu clustering and a potential new member of the NF kappa B  
 RT family within a 90 kilobase HLA class III segment."  
 RL Nat. Genet. 3:137-145(1993).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factor-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts."  
 RL Transplantation 61:1387-1392(1996).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwierzina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region."  
 RL Immunogenetics 42:315-322(1995).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface."  
 RL Cell 72:847-856(1993).  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,

RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373(1985).  
 [8]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=91086846; PubMed=1670639;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Rietmueller G., Weiss E.H.;  
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production.";  
 RL J. Exp. Med. 173:209-219(1991).  
 [9]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=91139175; PubMed=1671667;  
 RX Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNF gene.";  
 RL Immunogenetics 33:50-53(1991).  
 [10]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799(1994).  
 [11]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=95324911; PubMed=7601445;  
 RX Peelman L., Chardon P., Nunes M., Renard C., Geffroin C., Vaiman M.,  
 RA Van Zeveren A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The BAT1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family.";  
 RL Genomics 26:210-218(1995).  
 [12]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=20132445; PubMed=10668961;  
 RX Neville M.J., Campbell R.D.;  
 RA "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160(1997).  
 [13]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=98035883; PubMed=9367684;  
 RX Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidie U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing.";  
 RL Genomics 45:591-600(1997).  
 [14]  
 RA SEQUENCE FROM N.A.  
 RP MEDLINE=98149985; PubMed=9480751;  
 RX Shiina T., Taniya G., Oka A., Yamagata T., Yamagata N., Kikkawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 146-kilobase segment around the  
 RT IkbL and MICA genes at the centromeric end of the HLA class I  
 RT region.";  
 RL Genomics 47:372-382(1998).  
 [15]  
 RA SEQUENCE FROM N.A.  
 RP TISSUE=Lymphoid;  
 RC Blasanti R., Pessino A., Malaspina A.;  
 RA "NK-1 activating NK receptor.";  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 [16]  
 RA SEQUENCE FROM N.A.  
 RC TISSUE=Peripheral blood;  
 RA Sato M., Yabe T., Ohashi J., Tsuchiya N., Hanaoka K., Tokunaga K.,

RA Fuji T.;  
 RT "Identification of two novel single nucleotide polymorphisms in the  
 RT Nk30 gene in human natural killer cells.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 [17]  
 RA EMBL; AF031138; AAB86580.1; -;  
 DR EMBL; Y14768; CAA75065.1; -;  
 DR EMBL; AJ223153; CAB54004.1; -;  
 DR EMBL; AB055881; BAB78472.1; -;  
 DR GO; GO:0003793; P:defense/immunity protein activity; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR CO; GO:0006954; P:inflammatory response; NAS.  
 DR InterPro; IPR003599; IG.  
 DR Pfam; PF00047; IG; 1.  
 DR SMART; SMO0409; IG; 1.  
 DR PROSITE; PS0835; IG\_LIKE; 1.  
 KW Signal; Receptor.  
 FT SIGNAL 1 61 POTENTIAL.  
 FT CHAIN 62 190 IC7.  
 SQ SEQUENCE 190 AA; 20640 MW; 1FAC919E20A6B18A CRC64;  
 Query Match 89.4%; Score 76; DB 4; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 2.6e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 WVSQPPPIRTLEGS 14  
 Db 20 WVSQPPPIRTLEGS 33  
 RESULT 10  
 O14931  
 ID O14931 PRELIMINARY; PRT; 201 AA.  
 AC O14931;  
 DT 01-JAN-1998 (T-EMBLrel. 05, Created)  
 DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE I C7 precursor (NCR3 protein).  
 GN I C7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RX MEDLINE=96422187; PubMed=8924804;  
 RA Nalabolu S.R., Shukla H., Nallur G., Parimoo S., Weissman S.M.;  
 RT "Genes in a 220-kb region spanning the TNF cluster in human MHC.";  
 RL Genomics 31:215-222(1996).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Spleen;  
 RA Nalabolu S.R., Raghunathan A., Weissman S.M.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,  
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,  
 RA Lasky S., Hood L.;  
 RT "Sequence of the human major histocompatibility complex class III  
 RT region.";  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93272029; PubMed=8499947;  
 RA Iris F., Bougueret L., Prieur S., Caterina D., Primas G., Perrot V.,  
 RA Jurka J., Rodriguez-Tome P., Clavierie J., Cohen D., Dausset J.;  
 RT "Pense Alu clustering and a potential new member of the NF kappa B  
 RT family within a 90 kilobase HLA class III segment.";  
 RL Nat. Genet. 3:137-145(1993).  
 [5]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=96215741; PubMed=8629302;  
 RA Utans U., Quist W.C., McManus B.M., Wilson J.E., Arceci R.J.,  
 RA Wallace A.F., Russell M.E.;  
 RT "Allograft inflammatory factory-1. A cytokine-responsive macrophage  
 RT molecule expressed in transplanted human hearts.";  
 RL Transplantation 61:1387-1392 (1996).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96006565; PubMed=7590964;  
 RA Holzinger I., de Baey A., Messer G., Kick G., Zwiertzina H.,  
 RA Weiss E.H.;  
 RT "Cloning and genomic characterization of LST1: a new gene in the human  
 RT TNF region";  
 RL Immunogenetics 42:315-322 (1995).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93208891; PubMed=7916655;  
 RA Browning J.L., Ngan-ek A., Lawton P., DeMarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta: a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856 (1993).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86016093; PubMed=2995927;  
 RA Nedwin G.E., Naylor S.L., Sakaguchi A.Y., Smith D., Jarrett-Nedwin J.,  
 RA Pennica D., Goeddel D.V., Gray P.W.;  
 RT "Human lymphotoxin and tumor necrosis factor genes: structure,  
 RT homology and chromosomal localization.";  
 RL Nucleic Acids Res. 13:6361-6373 (1985).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91086846; PubMed=1670638;  
 RA Messer G., Spengler U., Jung M.C., Honold G., Bloemer K., Pape G.R.,  
 RA Riethmuller G., Weiss E.H.;  
 RT "Polymorphic structure of the tumor necrosis factor (TNF) locus: an  
 RT NcoI polymorphism in the first intron of the human TNF-beta gene  
 RT correlates with a variant amino acid in position 26 and a reduced  
 RT level of TNF-beta production.";  
 RL J. Exp. Med. 173:209-219 (1991).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53 (1991).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799 (1994).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vainan M.,  
 RA Van Zeven A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The B2M1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family.";  
 RL Genomics 26:210-218 (1995).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160 (1997).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;

RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,  
 RA Weiss E.H.;  
 RT "Complex expression pattern of the TNF region gene LST1 through  
 RT differential regulation, initiation, and alternative splicing.";  
 RL Genomics 45:591-600 (1997).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9814985; PubMed=9480751;  
 RA Shina T., Tamiya G., Oka A., Yamagata T., Yamagata N., Kikawa E.,  
 RA Goto K., Mizuki N., Watanabe K., Fukuzumi Y., Taguchi S., Sugawara C.,  
 RA Ono A., Chen L., Yamazaki M., Tashiro H., Ando S., Ikemura T.,  
 RA Kimura M., Inoko H.;  
 RT "Nucleotide sequencing analysis of the 145-kilobase segment around the  
 RT IkBL and MICA genes at the centromeric end of the HLA class I  
 RT region.";  
 RL Genomics 47:372-382 (1998).  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan R., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fehy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [17]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91139175; PubMed=1671667;  
 RA Abraham L.J., Du D.C., Zahedi K., Dawkins R.L., Whitehead A.S.;  
 RT "Haplotypic polymorphisms of the TNFB gene.";  
 RL Immunogenetics 33:50-53 (1991).  
 RN [18]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94362679; PubMed=8081366;  
 RA Albertella M.R., Campbell D.R.;  
 RT "Characterization of a novel gene in the human major  
 RT histocompatibility complex that encodes a potential new member of the  
 RT I kappa B family of proteins.";  
 RL Hum. Mol. Genet. 3:793-799 (1994).  
 RN [19]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95324911; PubMed=7601445;  
 RA Peelman L., Chardon P., Nunes M., Renard C., Geffrotin C., Vainan M.,  
 RA Van Zeven A., Coppieters W., Van de Weghe A., Bouquet Y., Choy W.,  
 RA Strominger J., Spies T.;  
 RT "The B2M1 gene in the MHC encodes an evolutionarily conserved putative  
 RT nuclear RNA helicase of the DEAD family.";  
 RL Genomics 26:210-218 (1995).  
 RN [20]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20132445; PubMed=10668961;  
 RA Neville M.J., Campbell R.D.;  
 RT "Alternative splicing of the LST-1 gene located in the major  
 RT histocompatibility complex on human chromosome 6.";  
 RL DNA Seq. 8:155-160 (1997).  
 RN [21]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98035883; PubMed=9367684;

Query Match 89.4%; Score 76; DB 4; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-05;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPPEIRITLLEGS 14

Db 20 WVSQPPPEIRITLLEGS 33

RESULT 11  
 Q8CG11

```

ID Q8CG11 PRELIMINARY; PRT; 192 AA.
AC Q8CG11; 2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 25, Last annotation update)
DE NK receptor 1c7 precursor.
GN 1c7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BN;
RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor 1c7.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430419; CAD23067.2; -.
DR EMBL; AJ430420; CAD23067.2; JOINED.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW SIGNAL; Receptor.
FT SIGNAL 1 19 POTENTIAL.
SQ SEQUENCE 192 AA; 20470 MW; 439AD7A3AFBE6DC0 CRC64;

Query Match 72.9%; Score 62; DB 11; Length 192;
Best Local Similarity 78.6%; Pred. No. 0.0078;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPEIRTLGSG 14
Db 20 WVSQPEIRTLGSGT 33

RESULT 12
Q8OWM8 PRELIMINARY; PRT; 192 AA.
ID Q8OWM8;
AC Q8OWM8;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE NKp30.
DE NKp30.
GN NKp30.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW;
RA Hsieh C.L., Obara H., Ali U., Rodriguez G., Nepomuceno R., Ogura Y.,
RA Martinez O.M., Krams S.M.;
RT "Identification, Cloning, and Characterization of a Novel Rat NK
RT Receptor, INKp30: a Molecule Expressed in Liver Allografts.";
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY273824; AAP13457.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 192 AA; 20498 MW; 7FDF58B245C52377 CRC64;

```

```

Query Match 72.9%; Score 62; DB 11; Length 192;
Best Local Similarity 78.6%; Pred. No. 0.0078;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPEIRTLGSG 14
Db 20 WVSQPEIRTLGSGT 33

RESULT 13
Q8CFD9 PRELIMINARY; PRT; 192 AA.
ID Q8CFD9;
AC Q8CFD9;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE 1c7 protein precursor.
GN 1c7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=NK cell;
RA Backman-Petersson E., Miller J.R., Hollyoake M., Aguado B.,
RA Butcher G.W.;
RT "Molecular characterization of the novel rat NK receptor 1c7.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430418; CAD23066.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF00047; IG; 1.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW SIGNAL.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 192 IC7 PROTEIN.
SQ SEQUENCE 192 AA; 20500 MW; 7FDD5AB252D239C7 CRC64;

Query Match 72.9%; Score 62; DB 11; Length 192;
Best Local Similarity 78.6%; Pred. No. 0.0078;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPEIRTLGSG 14
Db 20 WVSQPEIRTLGSGT 33

RESULT 14
Q86NN1 PRELIMINARY; PRT; 2201 AA.
ID Q86NN1;
AC Q86NN1;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE GH07949p (Fragment).
DE CG11936.
GN CG11936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dreesnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Munoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;

```

```

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT004503; A042667.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001025; BAH.
DR SMART; SM00439; BAH; 1.
FT NONTER
SQ SEQUENCE 2201 AA; 237100 MW; 10303133A8B4864B CRC64;

Query Match 57.6%; Score 49; DB 5; Length 2201;
Best Local Similarity 81.8%; Pred. No. 20;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 QPPEIRTLGDS 14
DB 340 QPPEIRTLGDS 350

RESULT 15
Q9VWCO PRELIMINARY; PRT; 2529 AA.
AC Q9VWCO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG32529 protein.
DE CG32529 OR CG11936 OR CG15619.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballwe R.M., Basu A., Baxendale J., Andrews-Bramnoch C., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brottier P.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bhandari D., Bolshakov S.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacle J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tuzy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Ciamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003513; AAP49026.2; -
DR FlyBase; FBGN0052529; CG32529.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001025; BAH.
DR SMART; SM00439; BAH; 1.
SQ SEQUENCE 2529 AA; 270675 MW; C93B1726EE7547D4 CRC64;

```

```

Query Match 57.6%; Score 49; DB 5; Length 2529;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 4 QPPEIRTLGDS 14
DB 668 QPPEIRTLGDS 678

```

```

Search completed: February 26, 2004, 12:13:50
Job time : 8.40584 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:00:57 ; Search time 9.31034 Seconds  
(without alignments)  
455.215 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPEIRTEUGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	15	4	Aae02774 Human Nkp
2	76	89.4	120	4	Aae02771 Human Nkp
3	76	89.4	135	5	Aae19109 Human Nkp
4	76	89.4	177	2	Aay06402 Human B-C
5	76	89.4	190	2	Aay06401 Human B-C
6	76	89.4	190	4	Aay02769 Human Nkp
7	76	89.4	201	2	Aay06403 Human B-C
8	76	89.4	369	5	Aae19110 Human Nkp
9	49	57.6	1289	4	Abb70840 Drosophila
10	45	52.9	164	4	Aau27977 Human con
11	45	52.9	301	3	Aag43485 Arabidops
12	45	52.9	353	3	Aag43484 Arabidops
13	45	52.9	353	5	Abg92033 Herbicida
14	45	52.9	380	3	Aag43483 Arabidops
15	44	51.8	185	4	Abg23939 Novel hum
16	44	51.8	688	3	Aay57393 Mouse ner
17	43	50.6	74	4	Aam89325 Human imm
18	43	50.6	276	4	Abg07919 Novel hum
19	43	50.6	514	5	Abj10466 XisF reco
20	42	49.4	63	5	Abp02197 Human ORF
21	42	49.4	134	3	Aag01336 Human sec
22	42	49.4	383	6	Abu39323 Protein e
23	42	49.4	451	2	Aay01520 Chicken C
24	42	49.4	459	6	Abp97695 Amino aci
25	42	49.4	462	5	Aau75313 Human Erg

## ALIGNMENTS

### RESULT 1

AAE02774

ID AAE02774 standard; peptide; 15 AA.

XX AAE02774;

XX AC AAE02774;

XX DT 06-AUG-2001 (first entry)

XX XX

DE Human Nkp30 receptor immunogenic peptide for antiserum production.

XX Human; Nkp30 receptor; natural killer cell; cytostatic; antimicrobial;  
 KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
 KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
 KW Therapy.

OS Homo sapiens.

XX WO200136630-A2.

XX PD 25-MAY-2001.

XX PF 15-NOV-2000; 2000WO-EP011697.

XX PR 15-NOV-1999; 99CA-02288307.

XX PA (INNA-) INNATE PHARMA SAS.

XX PA (UYGE-) UNIV GENOVA.

XX PI Moretta A, Bottino C, Biassoni R;

XX DR WPI; 2001-329221/34.

XX PT Novel compound, useful for detection and/or quantifying the presence of

XX FT NK cells, comprises the amino acid sequences of the Nkp30 molecule.

XX PS Claim 1; Page 33; 83pp; English.

XX CC The invention relates to human Nkp30 receptor and its corresponding cDNA

XX CC molecule which is involved in natural cytotoxicity mediated by natural

XX CC killer (NK) cells and antibodies that identify the same. Nkp30 receptor

XX CC is a member of immunoglobulin super family (Ig-SF). Nkp30 is selectively

XX CC expressed on the surface of human mature NK cells. Nkp30 and its cDNA are

XX CC useful for detecting and/or quantifying the presence of NK cells in a

XX CC biological sample. The invention also provide kits for detecting and/or

XX CC quantifying the presence of NK cells, for the selective removal of NK

XX CC cells from a biological sample, for the positive and selective

XX CC purification of NK cells from a biological sample and for the in vitro

Abu03476 Angiogene  
 Abus566 Lung canc  
 Adb75275 Prostate  
 Aay01521 Chicken c  
 Aau33622 Pseudomon  
 Abul5571 Protein e  
 Abp42291 Human ova  
 Abp33534 Human ORF  
 Aau80378 Human DOP  
 Abg02616 Novel hum  
 Aay49415 Human MTB  
 Aay59954 Human end  
 Aay03852 Murine le  
 Aay49414 Human T-b  
 Abm72545 Staphyloc  
 Abm67581 Drosophill  
 Aay57394 Human ner  
 Abg21153 Novel hum  
 Aao29646 Tribolium  
 Abb40612 Peptide #

26 42 49.4 462 6 ABU03476  
 27 42 49.4 462 6 ABUS566  
 28 42 49.4 462 7 ADB75275  
 29 42 49.4 478 2 AAY01521  
 30 42 49.4 1317 4 AAU33622  
 31 42 49.4 1317 6 ABUL5571  
 32 41.5 48.8 116 5 ABP42291  
 33 41 48.2 87 5 ABP33534  
 34 41 48.2 137 5 AAU80378  
 35 41 48.2 149 4 ABG02616  
 36 41 48.2 249 3 AAY49415  
 37 41 48.2 289 2 AAY59954  
 38 41 48.2 368 2 AAY03852  
 39 41 48.2 517 3 AAY49414  
 40 41 48.2 520 6 ABM72545  
 41 41 48.2 630 4 ABM67581  
 42 41 48.2 686 3 AAY57394  
 43 40.5 47.6 498 4 ABG21153  
 44 40.5 47.6 756 6 AAO29646  
 45 40 47.1 62 4 ABB40612

CC stimulation of NK cell cytotoxicity. The invention further provides a  
 CC pharmaceutical composition which is used as a drug for grafting  
 CC enhancement, graft versus host (GVH) inhibition, stimulation of graft  
 CC versus tumour (Gvt) and especially graft versus leukaemia (GvL), and for  
 CC the prevention, palliation and/or therapy of solid or liquid tumours,  
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or  
 CC microorganism, notably viral infection. NKp30 antibodies are useful for  
 CC identifying NKp30 natural ligands and allow assessment of the level of  
 CC surface NKp30 ligand expressed on an NK-susceptible target cell and the  
 CC comparison of this level to the standard physiological one. Hence NKp30  
 CC antibodies are useful in the diagnosis of tumours or of infection. The  
 CC present sequence is human NKp30 receptor immunogenic peptide fragment  
 CC which is used for NKp30 polyclonal antiserum production  
 XX  
 XX Sequence 15 AA;

Query Match 100.0%; Score 85; DB 4; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-07;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGSC 15  
 |||||  
 Db 1 WVSQPPEIRTLGSC 15

RESULT 2  
 AAE02771 ID AAE02771 standard; protein; 120 AA.  
 XX AAE02771;  
 AC AAE02771;  
 XX  
 DT 06-AUG-2001 (first entry)  
 XX  
 DE Human NKp30 receptor extracellular region sequence.  
 XX  
 KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;  
 KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
 KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
 KW therapy; extracellular region.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200136630-A2.  
 FN  
 XX  
 PD 25-MAY-2001.  
 XX  
 XX 15-NOV-2000; 2000WO-EP011697.  
 PF  
 XX 15-NOV-1999; 99CA-02288307.  
 PR  
 PR 15-NOV-1999; 99US-00440514.  
 XX  
 XX (INNA-) INNATE PHARMA SAS.  
 PA (UYGE-) UNIV GENOVA.  
 PA  
 XX Moretta A, Bottino C, Biassoni R;  
 XX  
 XX WPI; 2001-329221/34.  
 DR  
 XX  
 XX Novel compound, useful for detection and/or quantifying the presence of  
 PT NK cells, comprises the amino acid sequences of the NKp30 molecule.  
 PT  
 XX  
 XX Claim 1; Fig 7B; 83pp; English.  
 PS  
 XX

CC The invention relates to human NKp30 receptor and its corresponding cDNA  
 CC molecule which is involved in natural cytotoxicity mediated by natural  
 CC killer (NK) cells and antibodies that identify the same. NKp30 receptor  
 CC is a member of immunoglobulin super family (Ig-SF). NKp30 is selectively  
 CC expressed on the surface of human mature NK cells. NKp30 and its cDNA are  
 CC useful for detecting and/or quantifying the presence of NK cells in a  
 CC biological sample. The invention also provide kits for detecting and/or  
 CC quantifying the presence of NK cells, for the selective removal of NK  
 CC cells from a biological sample, for the positive and selective  
 CC purification of NK cells from a biological sample and for the in vitro

CC stimulation of NK cell cytotoxicity. The invention further provides a  
 CC pharmaceutical composition which is used as a drug for grafting  
 CC enhancement, graft versus host (GVH) inhibition, stimulation of graft  
 CC versus tumour (Gvt) and especially graft versus leukaemia (GvL), and for  
 CC the prevention, palliation and/or therapy of solid or liquid tumours,  
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or  
 CC microorganism, notably viral infection. NKp30 antibodies are useful for  
 CC identifying NKp30 natural ligands and allow assessment of the level of  
 CC surface NKp30 ligand expressed on an NK-susceptible target cell and the  
 CC comparison of this level to the standard physiological one. Hence NKp30  
 CC antibodies are useful in the diagnosis of tumours or of infection. The  
 CC present sequence is the extracellular region of human NKp30 receptor  
 XX  
 XX Sequence 120 AA;

Query Match 89.4%; Score 76; DB 4; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 0.00021;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGSG 14  
 |||||  
 Db 2 WVSQPPEIRTLGSG 15

RESULT 3  
 AAE19109 ID AAE19109 standard; protein; 135 AA.  
 XX AAE19109;  
 AC AAE19109;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human NKp30 protein.  
 XX  
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;  
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;  
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; NKp30.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200208287-A2.  
 FN  
 XX  
 PD 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-IL000664.  
 PF  
 XX 20-JUL-2000; 2000IL-00137419.  
 PR  
 XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 PA (UYNE ) UNIV BEN-GURION NEGEV.  
 PA  
 XX Mandelboim O, Porgador A;  
 PI  
 XX WPI; 2002-195870/25.  
 DR N-PSDB; AAD30466.  
 DR  
 XX  
 XX New targeting complex capable of targeting an active substance to a  
 PT target cell, comprising a target recognition segment and an active  
 PT segment, useful for treating pathologies associated with viral infections  
 PT or cancer.  
 XX  
 XX Example 1; Page 108; 113pp; English.  
 PS  
 XX

CC The invention relates to compositions and methods for the treatment and  
 CC detection of a variety of viral infections, by using complex agents  
 CC comprising the natural killer (NK) cells activating proteins, NKp46 and  
 CC NKp44 and functional fragments thereof, linked to therapeutic or imaging  
 CC agents. The complex is useful for treating pathologies associated with  
 CC viral infections (e.g. infections caused by influenza virus, HIV, Epstein  
 CC -Barr virus, cytomegalovirus, vaccinia virus, ECVV, MVM or herpes virus)  
 CC and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for  
 CC the imaging and monitoring of cancer. The complex may also be used to  
 CC detect the presence of abnormal cells in a sample. The antibodies can be



CC used to qualitatively or quantitatively detect the ligand for the  
CC complex. The present sequence is human MKP30 protein  
XX  
SQ Sequence 135 AA;

Query Match 89.4%; Score 76; DB 5; Length 135;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPERTLEGS 14  
|||  
Db 20 WVSQPPERTLEGS 33

RESULT 4  
AAV06402  
ID AAY06402 standard; protein; 177 AA.  
XX  
AC AAY06402;

XX 20-SEP-1999 (first entry)  
XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
XX  
XX MBOG; B-cell myelin oligodendrocyte glycoprotein; human;  
XX signal transduction; immunomodulator; antiinflammatory;  
XX autoimmune disease; inflammation; gene therapy; diagnosis.  
XX  
XX Homo sapiens.

Key	Location/Qualifiers
FT Peptide	1..12
FT Protein	/note= "leader peptide"
FT Modified-site	42
FT Modified-site	/note= "mature protein"
FT Modified-site	68
FT Modified-site	/note= "N-glycosylated"
FT Modified-site	121
FT Modified-site	/note= "N-glycosylated"
FT Domain	139..162
FT Peptide	/note= "transmembrane domain"
FT Peptide	166..177
FT Peptide	/note= "alternatively spliced C-terminal end"

WC9923867-A2.  
20-MAY-1999.

05-NOV-1998; 98WO-US023826.  
07-NOV-1997; 97US-0064761P.

(BIOJ ) BIOGEN INC.  
Browning J;

WPI; 1999-418423/35.  
N-PSDB; AAX59348.

Novel B-cell myelin oligodendrocyte glycoproteins.  
Claim 2; Page 43; 43pp; English.

This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of

CC producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)

SQ Sequence 177 AA;

Query Match 89.4%; Score 76; DB 2; Length 177;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPERTLEGS 14  
|||  
Db 20 WVSQPPERTLEGS 33

RESULT 5  
AAV06401  
ID AAY06401 standard; protein; 190 AA.  
XX  
AC AAY06401;

XX 20-SEP-1999 (first entry)  
XX Human B-cell myelin oligodendrocyte glycoprotein BMOG.  
XX  
XX MBOG; B-cell myelin oligodendrocyte glycoprotein; human;  
XX signal transduction; immunomodulator; antiinflammatory;  
XX autoimmune disease; inflammation; gene therapy; diagnosis.  
XX  
XX Homo sapiens.

Key	Location/Qualifiers
FT Peptide	1..12
FT Protein	/note= "leader peptide"
FT Modified-site	42
FT Modified-site	/note= "mature protein"
FT Modified-site	68
FT Modified-site	/note= "N-glycosylated"
FT Modified-site	121
FT Modified-site	/note= "N-glycosylated"
FT Domain	139..162
FT Peptide	/note= "transmembrane domain"
FT Peptide	166..190
FT Peptide	/note= "alternatively spliced C-terminal end"

WC9923867-A2.  
20-MAY-1999.

05-NOV-1998; 98WO-US023826.  
07-NOV-1997; 97US-0064761P.

(BIOJ ) BIOGEN INC.  
Browning J;

WPI; 1999-418423/35.  
N-PSDB; AAX59347.

Novel B-cell myelin oligodendrocyte glycoproteins.  
Claim 2; Page 42; 43pp; English.

This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AAY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph



CC nodes and in germinal centre B cells. It may have immunoregulatory  
 CC functions, and soluble or chimeric fusion proteins of BMOG may be used to  
 CC regulate the immune system in autoimmune or inflammatory disease. Vectors  
 CC comprising BMOG, prokaryotic and eukaryotic host cells, and a method of  
 CC producing BMOG using these transformed host cells are also provided. BMOG  
 CC polypeptides can be used for modulating the immune system of a subject or  
 CC to inhibit signal transduction in a cell expressing BMOG by contacting it  
 CC with a soluble BMOG protein. The nucleic acid can be used for gene  
 CC therapy. The protein can also be used to target a toxin, imaging agent or  
 CC radionuclide to a cell expressing BMOG. (All claimed)

XX SQ Sequence 190 AA;

Query Match 89.4%; Score 76; DB 2; Length 190;  
 Best Local Similarity 100.0%; Pred. No. 0.00033;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPPPIRTLEGS 14  
 |||||  
 Db 20 WVSQPPPIRTLEGS 33

#### RESULT 6

AAE02769  
 ID AAE02769 standard; protein; 190 AA.

XX AC AAE02769;

XX DT 06-AUG-2001 (first entry)

XX DE Human NKp30 receptor.

XX KW Human; NKp30 receptor; natural killer cell; cytostatic; antimicrobial;  
 KW melanoma; hepatocarcinoma; lung adenocarcinoma; viral infection; tumour;  
 KW immunosuppressant; antiviral; drug; grafting enhancement; leukaemia;  
 KW therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..18 /label= Signal\_peptide

FT Protein 19..190

FT Region 19..138 /label= Mature\_NKp30\_receptor\_protein

FT FT /label= Extracellular\_region  
 FT /note= "Forms an immunoglobulin (Ig) V-like domain"

FT Modified-site 42 /note= "N-glycosylation site"

FT Modified-site 121

FT Region 139..157 /note= "N-glycosylation site"

FT Region 158..190 /label= Transmembrane\_region

FT FT /label= Intracellular\_region

XX WO200136630-A2.

XX PD 25-MAY-2001.

XX PF 15-NOV-2000; 2000WO-EP011697.

XX PR 15-NOV-1999; 99CA-02288307.

XX PR 15-NOV-1999; 99US-00440514.

XX PA (INNA-) INNATE PHARMA SAS.

XX PA (UYGE-) UNIV GENOVA.

XX PI Moretta A, Bottino C, Biassoni R;

XX DR WPI; 2001-329221/34.

XX DR N-PSDB; AAD06564.

XX

FT Novel compound, useful for detection and/or quantifying the presence of  
 FT NK cells, comprises the amino acid sequences of the NKp30 molecule.

XX Claim 1; Fig 7B; 83pp; English.

XX CC The invention relates to human NKp30 receptor and its corresponding cDNA  
 CC molecule which is involved in natural cytotoxicity mediated by natural  
 CC killer (NK) cells and antibodies that identify the same. NKp30 receptor  
 CC is a member of immunoglobulin super family (Ig-SF). NKp30 is selectively  
 CC expressed on the surface of human mature NK cells. NKp30 and its cDNA are  
 CC useful for detecting and/or quantifying the presence of NK cells in a  
 CC biological sample. The invention also provide kits for detecting and/or  
 CC quantifying the presence of NK cells, for the selective removal of NK  
 CC cells from a biological sample, for the positive and selective  
 CC purification of NK cells from a biological sample and for the in vitro  
 CC stimulation of NK cell cytotoxicity. The invention further provides a  
 CC pharmaceutical composition which is used as a drug for grafting  
 CC enhancement, graft versus host (GVH) inhibition, stimulation of graft  
 CC versus tumour (GVT) and especially graft versus leukaemia (GVL), and for  
 CC the prevention, palliation and/or therapy of solid or liquid tumours,  
 CC such as melanoma, hepatocarcinoma and lung adenocarcinoma, and/or  
 CC microorganism, notably viral infection. NKp30 antibodies are useful for  
 CC identifying NKp30 natural ligands and allow assessment of the level of  
 CC surface NKp30 ligand expressed on an NK-susceptible target cell and the  
 CC comparison of this level to the standard physiological one. Hence NKp30  
 CC antibodies are useful in the diagnosis of tumours or of infection. The  
 CC present sequence is human NKp30 receptor

XX SQ Sequence 190 AA;

Query Match 89.4%; Score 76; DB 4; Length 190;

Best Local Similarity 100.0%; Pred. No. 0.00033;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WVSQPPPIRTLEGS 14

|||||

Db 20 WVSQPPPIRTLEGS 33

#### RESULT 7

AAE06403

ID AAY06403 standard; protein; 201 AA.

XX AC AAY06403;

XX DT 20-SEP-1999 (first entry)

XX DE Human B-cell myelin oligodendrocyte glycoprotein BMOG.

XX KW BMOG; B-cell myelin oligodendrocyte glycoprotein; human;  
 KW signal transduction; immunomodulator; antiinflammatory;  
 KW autoimmune disease; inflammation; gene therapy; diagnosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..12 /note= "leader peptide"

FT Protein 13..201 /note= "mature protein"

FT Modified-site 42 /note= "N-glycosylated"

FT Modified-site 68 /note= "N-glycosylated"

FT Modified-site 121 /note= "N-glycosylated"

FT Domain 139..162 /note= "transmembrane domain"

FT Peptide 166..201 /note= "alternatively spliced C-terminal end"

XX WO9923867-A2.

PD 20-MAY-1999.  
 XX  
 PF 05-NOV-1998; 98WO-US023826.  
 XX  
 PR 07-NOV-1997; 97US-0064761P.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Browning J;  
 XX  
 XX WPI; 1999-418423/35.  
 DR N-PSDB; AAXS9349.  
 XX  
 XX Novel B-cell myelin oligodendrocyte glycoproteins.  
 PT  
 XX  
 XX Claim 2; Page 43; 43pp; English.  
 PS  
 XX This sequence represents human BMOG, a novel member of the B cell myelin oligodendrocyte glycoprotein family that is expressed by germinal centre B cells. 3 C-terminal splice variants (see AY06401-03) of BMOG were identified. The protein is present primarily in the spleen, in lymph nodes and in germinal centre B cells. It may have immunoregulatory functions, and soluble or chimeric fusion proteins of BMOG may be used to regulate the immune system in autoimmune or inflammatory disease. Vectors comprising BMOG, prokaryotic and eukaryotic host cells, and a method of producing BMOG using these transformed host cells are also provided. BMOG polypeptides can be used for modulating the immune system of a subject or to inhibit signal transduction in a cell expressing BMOG by contacting it with a soluble BMOG protein. The nucleic acid can be used for gene therapy. The protein can also be used to target a toxin, imaging agent or radionuclide to a cell expressing BMOG. (All claimed)  
 XX  
 XX Sequence 201 AA;  
 SQ

Query Match 89.4%; Score 76; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 0.00035;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGGS 14  
 |||||  
 DB 20 WVSQPPEIRTLGGS 33

RESULT 8  
 AAE19110  
 ID AAE19110 standard; protein; 369 AA.  
 XX  
 AC AAE19110;  
 XX  
 DT 29-AUG-2003 (revised)  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Human NKp30-IgG fusion protein.  
 XX  
 KW Human; natural killer cell activating protein; NKp46; therapy; virucide;  
 KW viral infection; natural killer cell; NK; NKp44; imaging agent; cancer;  
 KW detection; carcinoma; melanoma; lymphoma; sarcoma; cytostatic; IgG;  
 KW immunoglobulin G; fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Chimeric.  
 XX  
 XX Key Location/Qualifiers  
 FH Region 1..135  
 FT /note= "Human NKp30"  
 FT Region 136..369  
 FT /note= "Human IgG"  
 XX  
 XX WO200208287-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-IL000664.  
 PF

XX 20-JUL-2000; 2000IL-00137419.  
 PR  
 XX (YISS ) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.  
 PA (UYNE ) UNIV BEN-GURION NEGEV.  
 XX  
 XX Mandelboim O, Forgador A;  
 PI  
 XX WPI; 2002-195870/25.  
 DR N-PSDB; AAD30467.  
 XX  
 XX New targeting complex capable of targeting an active substance to a target cell, comprising a target recognition segment and an active segment, useful for treating pathologies associated with viral infections or cancer.  
 PT  
 XX  
 XX Example 1; Page 108-110; 113pp; English.  
 PS  
 XX The invention relates to compositions and methods for the treatment and detection of a variety of viral infections, by using complex agents comprising the natural killer (NK) cells activating proteins, NKp46 and NKp44 and functional fragments thereof, linked to therapeutic or imaging agents. The complex is useful for treating pathologies associated with viral infections (e.g. infections caused by influenza virus, HIV, Epstein-Barr virus, cytomegalovirus, vaccinia virus, ECV, MVM or herpes virus) and cancer (e.g. carcinomas, melanomas, lymphomas and sarcomas), and for the imaging and monitoring of cancer. The complex may also be used to detect the presence of abnormal cells in a sample. The antibodies can be used to qualitatively or quantitatively detect the ligand for the complex. The present sequence is human NKp30- immunoglobulin G (IgG) FC region fusion protein. (Updated on 29-AUG-2003 to standardise OS field)  
 CC  
 XX Sequence 369 AA;  
 SQ

Query Match 89.4%; Score 76; DB 5; Length 369;  
 Best Local Similarity 100.0%; Pred. No. 0.00065;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPPEIRTLGGS 14  
 |||||  
 DB 20 WVSQPPEIRTLGGS 33

RESULT 9  
 ABB70840  
 ID ABB70840 standard; protein; 1289 AA.  
 XX  
 AC ABB70840;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 39312.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 XX WO200171042-A2.  
 XX  
 XX 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX Venter JC, Adams M, Li PWD, Myers EW;  
 PI WPI; 2001-656860/75.  
 DR N-PSDB; ABL14943.  
 DR

XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 XX Disclosure; SEQ ID NO 39312; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signaling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB016176-AB030511), expressed DNA  
 CC sequences (AB01840-AB018175) and the encoded proteins (AB057737-  
 CC AB072072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1289 AA;  
 Query Match 57.6%; Score 49; DB 4; Length 1289;  
 Best Local Similarity 81.8%; Pred. No. 57;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 QPPEIRTELEGS 14  
 ||| ||| ||| |||  
 Db 668 QPPEIRTELEGS 678  
 RESULT 10  
 AAU27977  
 ID AAU27977 standard; protein; 164 AA.  
 XX  
 AC AAU27977;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human contig polypeptide sequence #130.  
 XX  
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;  
 KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;  
 KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;  
 KW nervous system disorder; inflammatory disorder; cell differentiation;  
 KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;  
 KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;  
 KW cytoskeletal; antirheumatic; antiarthritic; vulnary; antinflammatory;  
 KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;  
 KW neuroprotective; osteopathic; antidiabetic; antiaesthetic; antiallergic;  
 KW immunostimulant; analgesic; gene therapy.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200164834-A2.  
 XX  
 XX 07-SEP-2001.  
 XX  
 XX 26-FEB-2001; 2001WO-US004926.  
 XX  
 XX 28-FEB-2000; 2000US-00515126.  
 PR 18-MAY-2000; 2000US-00577409.  
 PR 17-JUN-2000; 2000US-00597707.  
 PR 14-JUL-2000; 2000US-00616807.  
 PR 19-SEP-2000; 2000US-00664641.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F, Xu C;  
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;  
 PI Drmanac R;  
 XX  
 DR WPI; 2001-589862/66.  
 DR N-PDDB; AAG44877.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries  
 PT prepared from various human tissues, for diagnosis, treatment of cancer,  
 PT neurological, inflammatory disorders and for use in arrays for detection.  
 XX  
 XX Claim 10; Page 141-142; 153pp; English.  
 XX  
 CC Sequences AAU27676-AAU28019 represent full-length polypeptides and contig  
 CC polypeptides of the invention. The proteins and their associated DNA  
 CC sequences are useful for the treatment, diagnosis and prevention of  
 CC various types of disorder in a mammalian subject such as a human, dog,  
 CC monkey, mouse, hamster or rat. The disorders include cancers such as  
 CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as  
 CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,  
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system  
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's  
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and  
 CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's  
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory  
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,  
 CC cell proliferation, cell differentiation, stem cell growth factor,  
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells  
 CC in culture to give rise to neuroepithelial cells that can be used to  
 CC augment or replace cells damaged by illness, accidental damage or genetic  
 CC disorders. The sequences may also be used for regeneration of bone,  
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.  
 CC Note: Some sequences for this patent did not form part of the printed  
 CC specification, but were obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 164 AA;  
 Query Match 52.9%; Score 45; DB 4; Length 164;  
 Best Local Similarity 50.0%; Pred. No. 31;  
 Matches 9; Conservative 4; Mismatches 1; Indels 4; Gaps 1;  
 QY 1 WVSQPP-----EIRTELEGS 14  
 ||| ||| ||| ||| |||  
 Db 4 WVAQPPAVTADLQFLQLEGS 21  
 RESULT 11  
 AAG43485  
 ID AAG43485 standard; protein; 301 AA.  
 XX  
 AC AAG43485;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 54358.  
 XX  
 OS Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 XX 06-SEP-2000.  
 XX  
 XX 25-FEB-2000; 2000EP-00301439.  
 XX  
 XX 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125789P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.

PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	21-APR-1999;	99US-0130049P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131144P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145089P.
PR	07-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	11-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	21-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149829P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139751P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139837P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142134P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144335P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161320P.  
 PR 28-OCT-1999; 99US-0161322P.  
 PR 28-OCT-1999; 99US-0161393P.  
 PR 28-OCT-1999; 99US-0162142P.

Query Match 52.9%; Score 45; DB 3; Length 301;  
 Best Local Similarity 72.7%; Pred. No. 57;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 WVSQPEIRTL 11  
 Db 100 WVSQPKIREL 110

## RESULT 12

AAG43484  
 ID AAG43484 standard; protein; 353 AA.

XX AAG43484;

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 54357.

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000BP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136332P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.  
 PR 10-JUN-1999; 99US-0138540P.  
 PR 10-JUN-1999; 99US-0138847P.  
 PR 14-JUN-1999; 99US-0139119P.  
 PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139492P.  
 PR 18-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139763P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140495P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142820P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 20-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 21-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.

```
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145119P.
PR 23-JUL-1999; 99US-0145114P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 10-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 16-AUG-1999; 99US-0148684P.
PR 17-AUG-1999; 99US-0149368P.
PR 18-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154033P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159291P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.

PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 52.9%; Score 45; DB 3; Length 353;
Best Local Similarity 72.7%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPEIRTL 11
Db 152 WPSQPKIREL 162

RESULT 13
ABB92033
ID ABB92033 standard; protein; 353 AA.
AC ABB92033;
DT 31-MAY-2002 (first entry)
DE Herbicidally active polypeptide SEQ ID NO 1244.
KW Herbicidal; plant; agriculture; herbicide.
OS Arabidopsis thaliana.
PN WO200210210-A2.
PD 07-FEB-2002.
PF 28-AUG-2001; 2001WO-EP009892.
PR 28-AUG-2001; 2001WO-EP009892.
PA (FARB ) BAYER AG.
PI Tietjen X, Weidler M;
DR WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.
XX Claim 5; SEQ ID NO 1244; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins (ABB90790-ABB94016)
XX for herbicidally active compounds, comprising aligning and comparing
XX nucleic acid or amino acid sequences from plant with nucleic acid or
XX amino acid sequences from non-plant organisms using suitable search
XX parameters, where plant sequences having an E-value greater by a factor
XX of 3 than the E-value of most similar non-plant sequences are selected.
XX The polypeptides or nucleic acids encoding them are useful for
XX identifying modulators. The identified modulators are useful as
XX herbicides
XX Sequence 353 AA;
```

Query Match 52.9%; Score 45; DB 5; Length 353;  
Best Local Similarity 72.7%; Pred. No. 68;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPEIRTL 11  
Db 152 WVSQPKIREL 162

RESULT 14  
AAG43483  
ID AAG43483 standard; protein; 380 AA.  
AC AAG43483;  
XX  
DT 18-OCT-2000 (first entry)  
DE XX  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 54356.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
XX 05-MAR-1999; 99US-0123180P.  
XX 09-MAR-1999; 99US-0123548P.  
XX 23-MAR-1999; 99US-0125788P.  
XX 25-MAR-1999; 99US-0126264P.  
XX 29-MAR-1999; 99US-0126785P.  
XX 01-APR-1999; 99US-0127462P.  
XX 06-APR-1999; 99US-0128234P.  
XX 08-APR-1999; 99US-0128714P.  
XX 16-APR-1999; 99US-0129845P.  
XX 19-APR-1999; 99US-0130077P.  
XX 21-APR-1999; 99US-0130443P.  
XX 23-APR-1999; 99US-0130510P.  
XX 28-APR-1999; 99US-0130891P.  
XX 30-APR-1999; 99US-0132048P.  
XX 04-MAY-1999; 99US-0132407P.  
XX 05-MAY-1999; 99US-0132484P.  
XX 06-MAY-1999; 99US-0132486P.  
XX 06-MAY-1999; 99US-0132487P.  
XX 07-MAY-1999; 99US-0132863P.  
XX 11-MAY-1999; 99US-0132863P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134218P.  
XX 14-MAY-1999; 99US-0134221P.  
XX 14-MAY-1999; 99US-0134370P.  
XX 18-MAY-1999; 99US-0134768P.  
XX 19-MAY-1999; 99US-0134941P.  
XX 20-MAY-1999; 99US-0135124P.  
XX 21-MAY-1999; 99US-0135353P.  
XX 24-MAY-1999; 99US-0135629P.  
XX 25-MAY-1999; 99US-0136021P.  
XX 27-MAY-1999; 99US-0136392P.  
XX 28-MAY-1999; 99US-0136782P.  
XX 01-JUN-1999; 99US-0137222P.  
XX 03-JUN-1999; 99US-0137528P.  
XX 04-JUN-1999; 99US-0137502P.  
XX 07-JUN-1999; 99US-0137724P.  
XX 08-JUN-1999; 99US-0138094P.  
XX 10-JUN-1999; 99US-0138540P.  
XX 10-JUN-1999; 99US-0138847P.  
XX 14-JUN-1999; 99US-0139119P.  
XX 16-JUN-1999; 99US-0139452P.  
XX 16-JUN-1999; 99US-0139453P.  
XX 17-JUN-1999; 99US-0139492P.  
XX 18-JUN-1999; 99US-0139454P.  
XX 18-JUN-1999; 99US-0139455P.  
XX 18-JUN-1999; 99US-0139456P.  
XX 18-JUN-1999; 99US-0139457P.  
XX 18-JUN-1999; 99US-0139458P.  
XX 18-JUN-1999; 99US-0139459P.  
XX 18-JUN-1999; 99US-0139460P.  
XX 18-JUN-1999; 99US-0139461P.  
XX 18-JUN-1999; 99US-0139462P.  
XX 18-JUN-1999; 99US-0139463P.  
XX 18-JUN-1999; 99US-0139750P.  
XX 18-JUN-1999; 99US-0139763P.  
XX 21-JUN-1999; 99US-0139817P.  
XX 22-JUN-1999; 99US-0139899P.  
XX 23-JUN-1999; 99US-0140353P.  
XX 23-JUN-1999; 99US-0140354P.  
XX 24-JUN-1999; 99US-0140695P.  
XX 28-JUN-1999; 99US-0140823P.  
XX 29-JUN-1999; 99US-0140921P.  
XX 30-JUN-1999; 99US-0141267P.  
XX 01-JUL-1999; 99US-0141842P.  
XX 02-JUL-1999; 99US-0142154P.  
XX 06-JUL-1999; 99US-0142055P.  
XX 08-JUL-1999; 99US-0142390P.  
XX 09-JUL-1999; 99US-0142803P.  
XX 12-JUL-1999; 99US-0142920P.  
XX 13-JUL-1999; 99US-0142977P.  
XX 14-JUL-1999; 99US-0143542P.  
XX 14-JUL-1999; 99US-0143624P.  
XX 15-JUL-1999; 99US-0144005P.  
XX 16-JUL-1999; 99US-0144085P.  
XX 16-JUL-1999; 99US-0144086P.  
XX 19-JUL-1999; 99US-0144325P.  
XX 19-JUL-1999; 99US-0144331P.  
XX 19-JUL-1999; 99US-0144332P.  
XX 19-JUL-1999; 99US-0144333P.  
XX 19-JUL-1999; 99US-0144334P.  
XX 19-JUL-1999; 99US-0144335P.  
XX 20-JUL-1999; 99US-0144352P.  
XX 20-JUL-1999; 99US-0144632P.  
XX 20-JUL-1999; 99US-0144884P.  
XX 21-JUL-1999; 99US-0144814P.  
XX 21-JUL-1999; 99US-0145086P.  
XX 21-JUL-1999; 99US-0145088P.  
XX 22-JUL-1999; 99US-0145085P.  
XX 22-JUL-1999; 99US-0145087P.  
XX 22-JUL-1999; 99US-0145089P.  
XX 22-JUL-1999; 99US-0145192P.  
XX 23-JUL-1999; 99US-0145192P.  
XX 23-JUL-1999; 99US-0145145P.  
XX 23-JUL-1999; 99US-0145218P.  
XX 23-JUL-1999; 99US-0145224P.  
XX 26-JUL-1999; 99US-0145276P.  
XX 27-JUL-1999; 99US-0145913P.  
XX 27-JUL-1999; 99US-0145918P.  
XX 27-JUL-1999; 99US-0145919P.  
XX 28-JUL-1999; 99US-0145919P.  
XX 02-AUG-1999; 99US-0145951P.  
XX 02-AUG-1999; 99US-0145951P.  
XX 02-AUG-1999; 99US-0146388P.  
XX 02-AUG-1999; 99US-0146389P.  
XX 03-AUG-1999; 99US-0147038P.  
XX 04-AUG-1999; 99US-0147204P.  
XX 04-AUG-1999; 99US-0147302P.  
XX 05-AUG-1999; 99US-0147192P.  
XX 06-AUG-1999; 99US-0147260P.  
XX 06-AUG-1999; 99US-0147303P.  
XX 06-AUG-1999; 99US-0147416P.  
XX 09-AUG-1999; 99US-0147493P.  
XX 09-AUG-1999; 99US-0147935P.

PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 16-AUG-1999; 99US-0148684P.  
 PR 17-AUG-1999; 99US-0149336P.  
 PR 18-AUG-1999; 99US-0149175P.  
 PR 20-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150556P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151085P.  
 PR 27-AUG-1999; 99US-0151086P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 13-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159337P.  
 PR 18-OCT-1999; 99US-0159638P.  
 PR 21-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 52.9%; Score 45; DB 3; Length 380;  
 Best Local Similarity 72.7%; Pred. No. 73;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 WVSQPEIRTL 11  
 |||||

Db 179 WSPQPKIREL 189

# RESULT 15

ABG233939  
 ID ABG233939 standard; protein; 185 AA.

XX AC ABG233939;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #23930.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS88126.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 54298; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have application in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic  
 CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 185 AA;

Query Match 51.8%; Score 44; DB 4; Length 185;  
 Best Local Similarity 75.0%; Pred. No. 50;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 WVSQPEI 8  
 |:|||||

Db 158 WISQPEEL 165



Thu Feb 26 12:38:29 2004

us-10-036-444-7.rag

Page 12

Search completed: February 26, 2004, 12:10:18  
Job time : 12.5103 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:13:58 ; Search time 5.2122 Seconds  
(without alignments)  
607.670 Million cell updates/sec

Title: US-10-036-444-7

Perfect score: 85

Sequence: 1 WVSQPPPIRTLEGSC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	85	100.0	15	13	US-10-036-444-7
2	76	89.4	120	13	US-10-036-444-7
3	76	89.4	190	13	US-10-036-444-7
4	48	56.5	1517	14	US-10-156-761-14713
5	45.5	53.5	83	14	US-10-029-386-32542
6	43	50.6	514	15	US-10-014-099P-59
7	42	49.4	451	9	US-09-902-772-2
8	42	49.4	462	14	US-10-205-823-99
9	42	49.4	462	14	US-10-021-660-95
10	42	49.4	478	9	US-09-902-772-4
11	42	49.4	1317	9	US-09-815-242-5118
12	41.5	48.8	116	15	US-10-264-049-3423
13	41	48.2	87	11	US-09-864-408A-5014
14	41	48.2	137	10	US-09-947-063-2
15	41	48.2	137	10	US-09-947-063-5

16 41 48.2 339 15 US-10-369-493-16318  
17 41 48.2 368 14 US-10-166-183-6  
18 40 47.1 62 9 US-09-864-761-44601  
19 40 47.1 84 9 US-09-135-238B-4  
20 40 47.1 196 9 US-09-741-669-368  
21 40 47.1 196 9 US-09-912-020-376  
22 40 47.1 273 15 US-10-369-493-3987  
23 40 47.1 341 14 US-10-233-448-6  
24 40 47.1 547 15 US-10-080-334-276  
25 40 47.1 869 14 US-10-238-075-957  
26 39.5 46.5 1598 10 US-09-863-776-59  
27 39.5 46.5 1694 10 US-09-863-776-57  
28 39.5 46.5 1694 14 US-10-203-708-36  
29 39.5 46.5 1700 10 US-09-863-776-24  
30 39.5 46.5 1709 9 US-09-870-759-51  
31 39.5 46.5 1709 10 US-09-751-708A-51  
32 39.5 46.5 1709 10 US-09-863-776-58  
33 39.5 46.5 1709 10 US-09-863-776-60  
34 39.5 46.5 1709 14 US-10-203-708-35  
35 39 45.9 118 9 US-09-764-868-997  
36 39 45.9 170 14 US-10-305-555-36  
37 39 45.9 173 14 US-10-305-555-30  
38 39 45.9 200 9 US-09-908-193-6  
39 39 45.9 230 9 US-09-729-674-186  
40 39 45.9 307 9 US-09-816-087-2  
41 39 45.9 307 14 US-10-266-643-2  
42 39 45.9 307 14 US-10-305-555-2  
43 39 45.9 321 9 US-09-920-068A-2  
44 39 45.9 321 11 US-09-801-944B-268  
45 39 45.9 321 14 US-10-188-405-2

#### ALIGNMENTS

#### RESULT 1

US-10-036-444-7  
; Sequence 7, Application US/10036444  
; Publication No. US20020142445A1  
; GENERAL INFORMATION:  
; APPLICANT: INNATE PHARMA S.A.S.  
; APPLICANT: UNIVERSITA DI GENOVA  
; TITLE OF INVENTION: "NO. US20020142445A1 triggering receptor involved in natural cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same"  
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and antibodies that identify the same"  
; FILE REFERENCE: SEQ-PR-1060  
; CURRENT APPLICATION NUMBER: US/10/036,444  
; PRIOR FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: 09/440,514  
; PRIOR FILING DATE: 1999-11-15  
; PRIOR APPLICATION NUMBER: 09/456,199  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived from natural sequence, useful for antiserum  
; OTHER INFORMATION: production  
US-10-036-444-7

Query Match 100.0%; Score 85; DB 13; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.1e-06;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WVSQPPPIRTLEGSC 15

Db 1 WVSQPPPIRTLEGSC 15

```
RESULT 2
US-10-036-444-4
; Sequence 4, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-4
Query Match      89.4%; Score 76; DB 13; Length 120;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 WVSQPPPEIRTELEGS 14
Db 2 WVSQPPPEIRTELEGS 15

RESULT 3
US-10-036-444-2
; Sequence 2, Application US/10036444
; Publication No. US20020142445A1
; GENERAL INFORMATION:
; APPLICANT: INNATE PHARMA S.A.S.
; APPLICANT: UNIVERSITA DI GENOVA
; TITLE OF INVENTION: "No. US20020142445A1el triggering receptor involved in natural
; TITLE OF INVENTION: cytotoxicity mediated by human Natural Killer cells and
; TITLE OF INVENTION: antibodies that identify the same"
; FILE REFERENCE: SEQ-FR-1060
; CURRENT APPLICATION NUMBER: US/10/036.444
; PRIOR FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: 09/440,514
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 09/456,199
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Human NK cell
US-10-036-444-2
Query Match      89.4%; Score 76; DB 13; Length 190;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 WVSQPPPEIRTELEGS 14
Db 20 WVSQPPPEIRTELEGS 33

RESULT 4
US-10-156-761-14713
; Sequence 14713, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: KUEHN, Ralf
```

```
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14713
; LENGTH: 1517
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14713
Query Match      56.5%; Score 48; DB 14; Length 1517;
Best Local Similarity 61.5%; Pred. No. 68;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Cy 1 WVSQPPPEIRTELEG 13
Db 944 WVSQPPPEIRTELEG 956

RESULT 5
US-10-029-386-32542
; Sequence 32542, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32542
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011311.11
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P81408, EVALUE 5.30e-02
US-10-029-386-32542
Query Match      53.5%; Score 45.5; DB 14; Length 83;
Best Local Similarity 60.0%; Pred. No. 9.9;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Cy 1 WVSQPPPEIRTELEGSC 15
Db 24 W-SRPPELRPWGSC 37

RESULT 6
US-10-014-099F-59
; Sequence 59, Application US/10014099F
; Publication No. US20040003420A1
; GENERAL INFORMATION:
; APPLICANT: KUEHN, Ralf
```

```

; APPLICANT: FELDER, Susanne
; APPLICANT: SCHWENK, Frieder
; APPLICANT: KUESTER-LUKS, Birgit
; APPLICANT: FAUST, Nicole
; TITLE OF INVENTION: Modified Recombinase
; CURRENT APPLICATION NUMBER: US/10/014,099F
; CURRENT FILING DATE: 2001-11-12
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Xisf recombinase
; US-10-014-099F-59

Query Match          50.8%; Score 43; DB 15; Length 514;
Best Local Similarity 53.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy      2 VSPQPEIRTELEGS 14
Db      414 VESPPEVKTILRAS 426

RESULT 7
US-09-902-772-2
; Sequence 2, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of
; FILE REFERENCE: Chugai Seiyaku Kabushiki Kaisha 5001
; CURRENT APPLICATION NUMBER: US/09/902,772
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US/08/878,177
; PRIOR FILING DATE: 1997-06-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 451
; TYPE: PRT
; ORGANISM: protein sequence from C-11 gene
; US-09-902-772-2

Query Match          49.4%; Score 42; DB 9; Length 451;
Best Local Similarity 46.7%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTELEGSC 15
Db      56 WLSQPPARVTIKMEC 70

RESULT 8
US-10-205-823-99
; Sequence 99, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ganavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-205-823-99

Query Match          49.4%; Score 42; DB 14; Length 462;
Best Local Similarity 46.7%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTELEGSC 15
Db      63 WLSQPPARVTIKMEC 77

RESULT 9
US-10-021-660-95
; Sequence 95, Application US/10021660
; Publication No. US20030152926A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
; TITLE OF INVENTION: Modulators
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-021-660-95

Query Match          49.4%; Score 42; DB 14; Length 462;
Best Local Similarity 46.7%; Pred. No. 1.9e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTELEGSC 15
Db      63 WLSQPPARVTIKMEC 77

RESULT 10
US-09-902-772-4
; Sequence 4, Application US/09902772
; Patent No. US20020164739A1
; GENERAL INFORMATION:
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et

```



```
Db      68 WVGPEVKSENCC 82
|| : ||: |
RESULT 14
US-09-947-063-2
; Sequence 2, Application US/09947063
; Publication No. US20030059775A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru et al.
; TITLE OF INVENTION: NO. US20030059775A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-112
; CURRENT APPLICATION NUMBER: US/09/947,063
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/229,990
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/229,988
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-063-2
Query Match      48.2%; Score 41; DB 10; Length 137;
Best Local Similarity 46.2%; Pred. No. 83;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY      1 WVSQPPERTLEG 13
Db      105 WVGKDPQVRTVLG 117
|| : ||: |
RESULT 15
US-09-947-063-5
; Sequence 5, Application US/09947063
; Publication No. US20030059775A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru et al.
; TITLE OF INVENTION: NO. US20030059775A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-112
; CURRENT APPLICATION NUMBER: US/09/947,063
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/229,990
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/229,988
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-947-063-5
Query Match      48.2%; Score 41; DB 10; Length 137;
Best Local Similarity 46.2%; Pred. No. 83;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY      1 WVSQPPERTLEG 13
Db      105 WVGKDPQVRTVLG 117
|| : ||: |
Search completed: February 26, 2004, 12:33:59
Job time : 7.2122 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2004, 12:06:03 ; Search time 2.90451 Seconds  
(without alignments)  
266.616 Million cell updates/sec

Title: US-10-036-444-7

Sequence: 1 WVSQPEIRTEGSC 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB pep:\*  
5: /cgn2\_6/ptodata/2/iaa/6C COMB pep:\*  
6: /cgn2\_6/ptodata/2/iaa/6D COMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	49.4	451	3	US-08-878-177-2
2	42	49.4	478	3	US-08-878-177-4
3	42	49.4	1128	4	US-09-252-991A-31032
4	41	48.2	249	3	US-09-189-760-6
5	41	48.2	249	3	US-09-188-811-6
6	41	48.2	249	3	US-09-514-422-6
7	41	48.2	289	4	US-09-673-395A-155
8	41	48.2	478	4	US-09-134-000C-6219
9	41	48.2	517	3	US-09-189-760-2
10	41	48.2	517	3	US-09-514-422-2
11	41	48.2	567	3	US-09-188-811-2
12	40	47.1	202	4	US-09-328-352-5586
13	40	47.1	344	4	US-09-252-991A-33357
14	39	45.9	146	4	US-09-252-991A-32438
15	39	45.9	650	4	US-09-489-039A-7678
16	38	44.7	353	4	US-09-252-991A-19842
17	38	44.7	409	1	US-08-190-802A-51
18	38	44.7	409	2	US-08-283-917-3
19	38	44.7	409	2	US-08-961-716-3
20	38	44.7	409	3	US-08-477-346-51
21	38	44.7	409	4	US-08-473-089-51
22	38	44.7	409	4	US-08-487-072A-51
23	38	44.7	410	2	US-08-283-917-9
24	38	44.7	410	2	US-08-961-716-9
25	38	44.7	414	4	US-09-543-681A-7732
26	38	44.7	414	4	US-09-543-681A-7982
27	38	44.7	675	3	US-09-171-878-1

28	38	44.7	947	2	US-08-887-518-2	Sequence 2, Appli
29	38	44.7	947	2	US-09-023-321-2	Sequence 2, Appli
30	38	44.7	947	2	US-09-032-475-2	Sequence 2, Appli
31	38	44.7	947	3	US-09-257-703-1	Sequence 1, Appli
32	38	44.7	947	4	US-09-871-889A-1	Sequence 1, Appli
33	38	44.7	2972	4	US-08-469-260A-387	Sequence 387, App
34	38	44.7	2972	4	US-08-488-446-387	Sequence 387, App
35	38	44.7	2972	4	US-08-467-344A-387	Sequence 387, App
36	37	43.5	107	4	US-09-345-236B-27	Sequence 27, Appli
37	37	43.5	256	4	US-09-483-039A-9057	Sequence 9057, Ap
38	37	43.5	263	5	PCT-US91-08177-13	Sequence 13, Appli
39	37	43.5	280	4	US-09-252-991A-20783	Sequence 20783, A
40	37	43.5	370	4	US-09-489-039A-8738	Sequence 8738, Ap
41	37	43.5	449	2	US-08-819-458A-16	Sequence 16, Appli
42	37	43.5	490	4	US-09-056-285A-10	Sequence 10, Appli
43	37	43.5	496	2	US-08-463-418-2	Sequence 2, Appli
44	37	43.5	577	4	US-09-252-991A-19546	Sequence 19546, A
45	37	43.5	602	2	US-08-413-652-6	Sequence 6, Appli

## ALIGNMENTS

### RESULT 1

US-08-878-177-2 Application US/08878177  
; Sequence 2, Application US/08878177  
; Patent No. 6294354  
; GENERAL INFORMATION:  
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et  
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
; TITLE OF INVENTION: the Proteins  
; FILE REFERENCE: chugai seiyaku kabushiki Kaisha 5001  
; CURRENT APPLICATION NUMBER: US/08/878,177  
; CURRENT FILING DATE: 1997-06-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 451  
; TYPE: PRT  
; ORGANISM: protein sequence from C-11 gene  
US-08-878-177-2

Query Match 49.4%; Score 42; DB 3; Length 451;  
Best Local Similarity 46.7%; Pred. No. 56;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY	1	WVSQPEIRTEGSC	15
DB	56	WLSQPPARVTIKMEC	70

### RESULT 2

US-08-878-177-4 Application US/08878177  
; Sequence 4, Application US/08878177  
; Patent No. 6294354  
; GENERAL INFORMATION:  
; APPLICANT: Chugai Seiyaku Kabushiki Kaisha, 5001, Iwamoto et  
; TITLE OF INVENTION: Cell Calcification Suppressing Proteins and Genes of  
; TITLE OF INVENTION: the Proteins  
; FILE REFERENCE: chugai seiyaku kabushiki Kaisha 5001  
; CURRENT APPLICATION NUMBER: US/08/878,177  
; CURRENT FILING DATE: 1997-06-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: protein sequence from c-erg gene  
US-08-878-177-4

Query Match 49.4%; Score 42; DB 3; Length 478;  
Best Local Similarity 46.7%; Pred. No. 70;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
OY 1 WVSQPPERTLESGC 15
|:|||||:|
Db 56 WLSQPPARVTKMEC 70

RESULT 3
US-09-252-991A-31032
; Sequence 31032, Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31032
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31032

Query Match 49.4%; Score 42; DB 4; Length 1128;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPPERTLEGS 14
|:|||||:|
Db 155 WAAVPAELQTOEGS 168

RESULT 4
US-09-189-760-6
; Sequence 6, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-189-760-6

Query Match 48.2%; Score 41; DB 3; Length 249;
Best Local Similarity 35.7%; Pred. No. 51;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPPERTLEGS 14
|:|||||:|
Db 181 WIETPPSIKSLDSN 194

RESULT 5
US-09-188-811-6
```

```
; Sequence 6, Application US/09188811
; Patent No. 6037148
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP
; CURRENT APPLICATION NUMBER: US/09/188,811
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-188-811-6

Query Match 48.2%; Score 41; DB 3; Length 249;
Best Local Similarity 35.7%; Pred. No. 51;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPPERTLEGS 14
|:|||||:|
Db 181 WIETPPSIKSLDSN 194

RESULT 6
US-09-514-422-6
; Sequence 6, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/514,422
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US/09/189,760
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/163,116
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/089,467
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: (PENDING)
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-514-422-6

Query Match 48.2%; Score 41; DB 3; Length 249;
Best Local Similarity 35.7%; Pred. No. 51;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 WVSQPPERTLEGS 14
|:|||||:|
Db 181 WIETPPSIKSLDSN 194

RESULT 7
US-09-673-395A-155
; Sequence 155, Application US/09673395A
; Patent No. 8620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
```



```
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 155
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-155

Query Match      48.2%; Score 41; DB 4; Length 289;
Best Local Similarity 46.7%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 15
Db      267 WQQPFAARSCYGLC 281

RESULT 8
US-09-134-000C-6219
; Sequence 6219, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6219
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-6219

Query Match      48.2%; Score 41; DB 4; Length 478;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 15
Db      290 WIENPFLLELVGTC 304

RESULT 9
US-09-189-760-2
; Sequence 2, Application US/09189760
; Patent No. 6031078
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/189,760
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; EARLIER APPLICATION NUMBER: 60/089,467
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: (PENDING)
; EARLIER FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-189-760-2

Query Match      48.2%; Score 41; DB 3; Length 517;
Best Local Similarity 35.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 14
Db      449 WIETPPSIKSLDSN 462

RESULT 10
US-09-514-422-2
; Sequence 2, Application US/09514422
; Patent No. 6291193
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP2
; CURRENT APPLICATION NUMBER: US/09/514,422
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US/09/189,760
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 09/163,116
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/089,467
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: (PENDING)
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-514-422-2

Query Match      48.2%; Score 41; DB 3; Length 517;
Best Local Similarity 35.7%; Pred. No. 1.1e+02;
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy      1 WVSQPPPEIRTLGSC 14
Db      449 WIETPPSIKSLDSN 462

RESULT 11
US-09-188-811-2
; Sequence 2, Application US/09188811
; Patent No. 6037148
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL MTBX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-046CP
; CURRENT APPLICATION NUMBER: US/09/188,811
; CURRENT FILING DATE: 1998-11-09
; EARLIER APPLICATION NUMBER: 09/163,116
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-188-811-2

Query Match      48.2%; Score 41; DB 3; Length 567;
Best Local Similarity 35.7%; Pred. No. 1.2e+02;
```

```
Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 WVSQPPEIRTEGSC 14
   ||| |||
Db 499 WISPPSIKSLDSN 512

RESULT 12
US-09-328-352-5586
; Sequence 5586, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5586
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5586

Query Match 47.1%; Score 40; DB 4; Length 202;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 WVSQPPEIRTEGSC 14
   ||| |||
Db 157 WLATPEIRKLGGA 170

RESULT 13
US-09-252-991A-31357
; Sequence 31357, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31357
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31357

Query Match 47.1%; Score 40; DB 4; Length 344;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 WVSQPPEIR 9
   ||| |||
Db 185 WASLPPEVR 193

RESULT 14
US-09-252-991A-32438
; Sequence 32438, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32438
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32438

Query Match 45.9%; Score 39; DB 4; Length 146;
Best Local Similarity 54.5%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 5 PPEIRTEGSC 15
   ||| |||
Db 72 PPVRSRSGSC 82

RESULT 15
US-09-489-039A-7678
; Sequence 7678, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7678
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7678

Query Match 45.9%; Score 39; DB 4; Length 650;
Best Local Similarity 63.6%; Pred. No. 3e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 5 PPEIRTEGSC 15
   ||| |||
Db 446 PTEVLTLEGFC 456
```

Search completed: February 26, 2004, 12:16:20  
Job time : 4.90451 secs